The Biometric Society

BIOMETRICS

FOUNDED BY THE BIOMETRICS SECTION OF THE AMERICAN STATISTICAL ASSOCIATION

TABLE OF CONTENTS

Number 2 June 1954 Volume 10

Material for *Biometrics* should be addressed to Miss Gertrude Cox, Institute of Statistics, Box 5457, Raleigh, North Carolina, except that authors residing in one of the following organized regions can expedite the handling of their papers by submitting them to the Assistant Editor for that region.

British Region: Dr. D. J. Finney, 6 Keble Road, Oxford, England; **Australasian Region**: Dr. E. A. Cornish, University of Adelaide, Adelaide, Australia; **French Region**: Dr. Georges Teissier, Faculte des Sciences de Paris, 1 rue V. Cousin, Paris, France.

Material for Queries should go to Professor G. W. Snedecor, Statistical Laboratory, Iowa State College, Ames, Iowa.

Articles to be considered for publication should be submitted in triplicate.

The Biometric Society is an international society devoted to the mathematical and statistical aspects of biology and welcomes to membership biologists, mathematicians, statisticians and others who are interested in its objectives. Through its regional organizations the Society sponsors regional and local meetings. National secretaries serve the interest of members in Denmark, the Netherlands, India, Germany, Japan, Sweden and Brazil and there are many members "at large". Dues in the Society for 1954 for residents of the Western Hemisphere are as follows: Full membership including subscription to *Biometrics* is $7.00. Members of the Biometrics Section of the American Statistical Association who subscribe to the journal through that organization may become members of The Biometric Society on the payment of $3.00 annual dues. For members in other parts of the world, full membership including subscription to *Biometrics* is $4.50, except that members who subscribe to the journal through the American Statistical Association pay annual dues of $1.75. Information concerning the Society can be obtained from the Secretary, The Biometric Society, Drawer 1106, New Haven 4, Connecticut, U.S.A.

Annual subscription rates to non-members are as follows: For American Statistical Association Members, $4.00; for subscribers, non-members of either American Statistical Association or The Biometric Society, $7.00. Subscriptions should be sent to the Managing Editor, *Biometrics*, P. O. Box 5457, Raleigh, North Carolina, U.S.A.

Entered as second-class matter at the Post Office at New Haven, Conn., under the Act of March 3, 1879. Additional entry at Richmond, Va. Business Office, 52 Hillhouse Ave., New Haven, Conn. *Biometrics* is published quarterly—in March, June, September and December.

FURTHER STUDIES ON THE SIGNIFICANCE OF FAMILY FACTORS FOR THE RESPONSE TO BCG VACCINATION.

THE DEVELOPMENT OF LOCAL VACCINATION LESIONS AND THEIR RELATION TO ALLERGY PRODUCTION.

SVEN NISSEN MEYER AND MICHAEL WEIS BENTZON

Tuberculosis Research Office, World Health Organization, Copenhagen, Denmark

Results presented in two previous papers (1,2) have shown that tuberculin allergy developing after BCG vaccination depends on the family membership of the vaccinated child. According to these results, the degree of BCG-induced allergy could be regarded as a sum of two variables, (1) a "family value" determined by the family membership of the child, and (2) a positive or negative deviation from this family value, which may modify the reaction observed in the individual child. This latter deviation may originate from various causes—from biological differences within sibling groups, from random errors in technique and dosage of vaccination and finally from errors made in the tests used for measuring allergy. However, in the following analysis it is convenient to let these latter causes be represented by a single variable.

The purpose of the present paper is, first, to demonstrate a similar influence of family factors on the production of the *local vaccination lesion*. Second, the manifestation of family factors in separate measurable effects of the vaccination—allergy and local lesion—suggests an investigation of their interrelation. The question arises whether the family factors appearing in the various types of responses are identical (i.e. actually express the same family property) and if not, whether they are correlated or uncorrelated. An approach is made to this problem, and some biological implications of the results are discussed.

1. MATERIAL

Details about material and testing technique have been given in the preceding papers, and only principal points will be repeated here.

The material was obtained from an investigation on BCG vaccination, conducted during the period November 1949-February 1950 among school children from a rural area in Denmark. Essentially all children were in the age span 7-14 years and 51% were boys. Only previously unvaccinated children, giving less than 6 mm induration to an intradermal Mantoux test with 10 TU*, were included in the study.

*1 TU (tuberculin unit) = 1/50000 mg ref. standard PPD or 0.01 mg international standard O.T. (0.1 cc of 1/10000 dilution).

Vaccination of these children was carried out with 39 samples of vaccine #869 from the State Serum Institute in Copenhagen, graduated with respect to dosage, age of vaccine, and temperature of storage. The same sample of vaccine was used to perform all vaccinations within any given school, each sample providing for from 1 to 4 schools.

The diameter of the local lesion developing at the site of vaccination was carefully measured 10 weeks after vaccination. Mantoux tests with 10 TU were given after the same period, and the transverse diameter of induration—recorded 3 or 4 days later (constant reading interval within each school)—was taken as a quantitative measure of the degree of post-vaccination allergy. This re-examination after 10 weeks comprised 84 schools with 1733 children belonging to 731 families, each with 2-5 vaccinated children.

Mantoux tests with 10 TU were carried out again one year after vaccination on 1085 children attending 86 schools and belonging to 485 families. Included in both retestings were 72 schools, 898 children and 401 families.

2. PRINCIPLES OF THE STATISTICAL ANALYSIS

The appropriate method for demonstrating familial differences in a response is the analysis of variance, used also in the two previous studies on tuberculin allergy. The Mantoux reactions were suitable for this analysis insofar as they were approximately normally distributed by size of induration. However, the design of the field investigation implied that several factors, such as use of the same vaccine ampule, uniform testing and reading conditions, easily could produce differences between the schools. Although there was no correlation between mean values and standard deviations, both characteristics showed a significant variation from school to school. It was necessary, therefore, to analyse each school separately for family differences. As sampling errors could be expected to influence the results obtained from the individual schools, a χ^2-test was finally applied to the distribution of all 84 variance ratios.

The sizes of vaccination lesions gave skewed distributions, and there was a distinct positive correlation between mean values and standard deviations—both characteristics increased with increasing strength of the vaccine. As illustrated in Figure 1 a-b, a logarithmical transformation of the sizes of vaccination lesions resulted in approximately normal distributions. The figures show probit diagrams for the measured size of vaccination lesions and for the logarithmically transformed sizes—the total of all 84 schools being divided in three major groups, each of which has been treated with vaccines of approximately

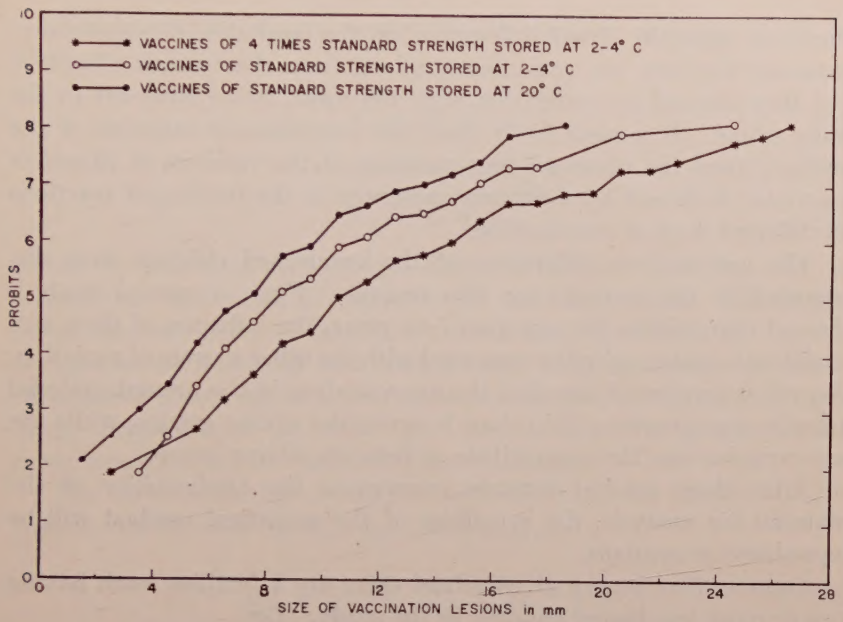

FIGURE 1a. PROBIT DIAGRAM OF SIZES OF VACCINATION LESIONS

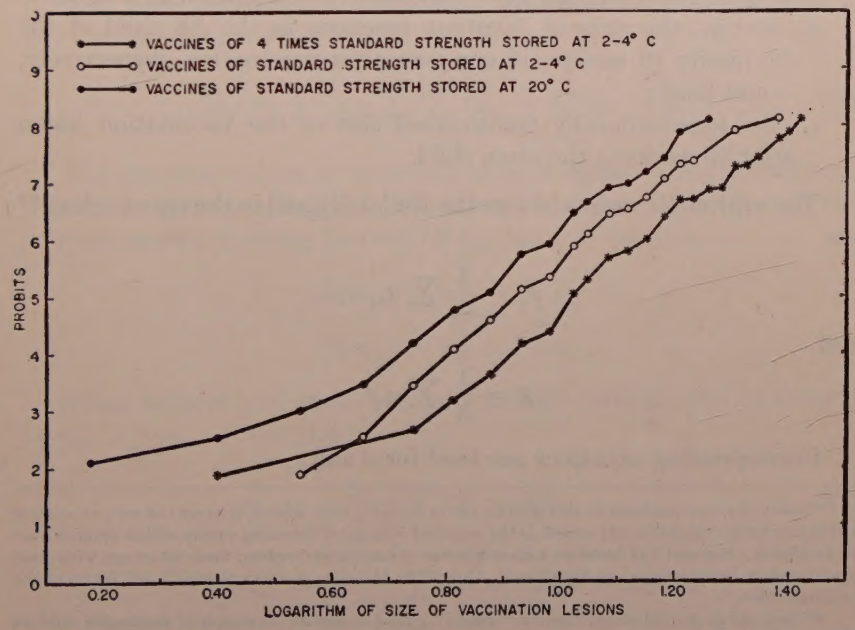

FIGURE 1b. PROBIT DIAGRAM OF THE LOGARITHMICALLY TRANSFORMED SIZES OF VACCINATION LESIONS

the same strength. Small differences, on the borderline of significance, remained between the variances obtained from the various schools*), but they showed no correlation with the mean value observed in the same school. It is most likely that this unsystematic variation of the variance (and the corresponding variation of the variance of Mantoux reactions) is caused by a different accuracy in the reading of reactions on different days of examination.

The age and sex differences of the vaccinated children were disregarded in the analysis for two reasons. First, a special analysis showed that, within the age span 7-14 years, the influence of these two variables is quite negligible compared with the other sources of variation. Second, it was found also that the age-variation in the present material actually was greater within than between the sibling groups, while the sex-variation was the same within as between sibling groups.

After these general remarks concerning the applicability of the material for analysis, the principles of the statistical method will be reproduced in symbols.

Suppose that in any given school there are k families, each having two or more vaccinated children in the study. Let

n_i denote the number of vaccinated children in the ith family,
$N = \sum n_i$ the total number of vaccinated children from all k families,
x_{ij} and y_{ij} the sizes of Mantoux reactions in the jth child of the ith family 10 weeks and one year after vaccination, respectively,
—and finally
z_{ij} the logarithmically transformed size of the vaccination lesion after 10 weeks in the same child.

The arithmetic means of x in the ith family and in the entire school** are

$$\bar{x}_i = \frac{1}{n_i} \sum x_{ij}$$

and

$$\bar{x} = \frac{1}{N} \sum n_i \bar{x}_i$$

Corresponding notations are used for y and z.

*Briefly the test consisted in establishing ratios $Q/\overline{s^2}$ for each school, Q being the sum of squares within families for the particular school, $\overline{s^2}$ the weighted average of the mean square within families over all 84 schools. Because $\overline{s^2}$ is based on a great number of degrees of freedom, these ratios can with good approximation be regarded as χ^2-distributed. Out of the 84 ratios, 9 were outside the 5% limits of the χ^2-distribution.

**Here and in the following, the term "school" is used to denote the sample of vaccinated children from families having at least two vaccinated children within the school, i.e. children having no vaccinated siblings in the school are excluded.

The hypothesis that is tested by the analysis of variance can be expressed as follows:

$$x_{ij} = \xi_i + u_{ij} \tag{1a}$$

$$y_{ij} = \eta_i + v_{ij} \tag{1b}$$

$$z_{ij} = \zeta_i + w_{ij} \tag{1c}$$

where the first term on the right side denotes the "family value", the second the "individual deviation" from the family value. (It may be noted that according to this model, the size of the tuberculin reaction is regarded as a *sum* of two components, while the measured size of vaccination lesion will be a *product* of two quantities, one depending on family properties of the child, the other on individual properties and experimental errors in vaccination and readings.)

The three estimates of variance obtained from the variation *within* families are denoted by m_{uu} , m_{vv} , and m_{ww} , i.e.:

$$m_{uu} = \frac{1}{N-k} \sum \sum (x_{ij} - \bar{x}_i)^2 \tag{2a}$$

etc.

The three estimates of covariance within families are denoted by m_{uv} , m_{uw} and m_{vw} , i.e.:

$$m_{uv} = \frac{1}{N-k} \sum \sum (x_{ij} - \bar{x}_i)(y_{ij} - \bar{y}_i) \tag{2b}$$

The expected values of these estimates of variances and covariances will be the corresponding population moments of the variables u, v and w, these moments being denoted by μ_{uu} , μ_{uv} . . . etc., i.e.:

$$E(m_{uu}) = \mu_{uu}$$

$$E(m_{uv}) = \mu_{uv} \ldots \text{etc.}$$

Mean squares and mean products *between* families will be denoted by $m_{\bar{x}\bar{x}} \ldots m_{\bar{x}\bar{y}} \ldots$ etc., i.e.:

$$m_{\bar{x}\bar{x}} = \frac{1}{k-1} \sum (\bar{x}_i - \bar{x})^2 n_i \tag{3a}$$

$$m_{\bar{x}\bar{y}} = \frac{1}{k-1} \sum (\bar{x}_i - \bar{x})(\bar{y}_i - \bar{y}) n_i \tag{3b}$$

and correspondingly for the other variables.

These estimates will have the following expected values:

$$E(m_{\bar{x}\bar{x}}) = \mu_{uu} + \mu_{\xi\xi}\frac{1}{k-1}\left(\sum n_i - \frac{\sum n_i^2}{\sum n_i}\right) \tag{4a}$$

$$E(m_{\bar{x}\bar{y}}) = \mu_{uv} + \mu_{\xi\eta}\frac{1}{k-1}\left(\sum n_i - \frac{\sum n_i^2}{\sum n_i}\right) \tag{4b}$$

etc., the family values ξ, η and ζ are here regarded as random variables with population moments denoted by $\mu_{\xi\xi}$, $\mu_{\xi\eta}$ . . . etc.

As the number of siblings per family shows little variation, the last term in the expressions $(4\ a\ -\ b)$ can with good approximation be replaced by $\mu_{\xi\xi}\cdot r$ and $\mu_{\xi\eta}\cdot r$, r being the average number of siblings per family. The variances and covariances of the family values in the population investigated can then be estimated by:

$$\mu_{\xi\xi} \approx (m_{\bar{x}\bar{x}} - m_{uu})\frac{1}{r} \tag{5}$$

etc.

Assuming that the significant variation of $m_{\bar{x}\bar{x}}$ and m_{uu} between the schools is due to differences in the sizes of experimental errors made on different days, we can in (5) replace $m_{\bar{x}\bar{x}}$ and m_{uu} by weighted averages obtained from all schools. The variances and covariances of the family values can then be estimated from a greater number of observations.

The next problem to be investigated is whether there is a correlation or even an exact functional relation between the three family values. The hypothesis of functional dependency, expressed analytically as follows

$$\xi = f(\theta), \qquad \eta = g(\theta), \qquad \zeta = h(\theta)$$

would mean that the three types of responses actually reflect the same basic family property θ. We shall test the special case of a linear relationship, i.e. the hypothesis:

$$\xi_i = \alpha_1 + \beta_1\theta_i \tag{6a}$$

$$\eta_i = \alpha_2 + \beta_2\theta_i \tag{6b}$$

$$\zeta_i = \alpha_3 + \beta_3\theta_i \tag{6c}$$

where α_n and β_n are constants which can be chosen so that $\bar{\theta} = 0$.

(It may be noted that equation (6c) gives an exponential relation between the measured size of vaccination lesion and the basic family variable).

Inserting (6 a − c) in (1 a − c) we get:

$$x_{ij} = \alpha_1 + \beta_1\theta_i + u_{ij} \tag{7a}$$

$$y_{ij} = \alpha_2 + \beta_2\theta_i + v_{ij} \tag{7b}$$

$$z_{ij} = \alpha_3 + \beta_3\theta_i + w_{ij} \tag{7c}$$

Assuming a normal distribution of u, v, w, it follows that the variable

$$t_{ij} = y_{ij} - \frac{\beta_2}{\beta_1}x_{ij} - \alpha_2 + \frac{\beta_2}{\beta_1}\alpha_1 = v_{ij} - \frac{\beta_2}{\beta_1}u_{ij} \tag{8}$$

has a normal distribution with a mean of zero and a variance of:

$$\sigma^2 = \mu_{vv} + \left(\frac{\beta_2}{\beta_1}\right)^2\mu_{uu} - 2\left(\frac{\beta_2}{\beta_1}\right)\mu_{uv} \tag{9}$$

The two sums of squares:

$$Q_1 = \sum n_i(\bar{t}_i - \bar{t})^2 = \sum n_i\left[(\bar{y}_i - \bar{y}) - \frac{\beta_2}{\beta_1}(\bar{x}_i - \bar{x})\right]^2 \tag{10}$$

and

$$Q_2 = \sum\sum(t_{ij} - \bar{t}_i)^2 = \sum\sum\left[(y_{ij} - \bar{y}_i) - \frac{\beta_2}{\beta_1}(x_{ij} - \bar{x}_i)\right]^2 \tag{11}$$

are stochastically independent and distributed as $\sigma^2\chi^2$ with $(k - 1)$ and $(N - k)$ degrees of freedom, respectively.

It follows from (2 a − b) and (3 a − b) that

$$Q_1 = (k - 1)\left[m_{\bar{y}\bar{y}} + \left(\frac{\beta_2}{\beta_1}\right)^2 m_{\bar{x}\bar{x}} - 2\left(\frac{\beta_2}{\beta_1}\right)m_{\bar{x}\bar{y}}\right] \tag{12}$$

and

$$Q_2 = (N - k)\left[m_{vv} + \left(\frac{\beta_2}{\beta_1}\right)^2 m_{uu} - 2\left(\frac{\beta_2}{\beta_1}\right)m_{uv}\right] \tag{13}$$

The ratio

$$F = \frac{(N - k)Q_1}{(k - 1)Q_2} = \frac{m_{\bar{y}\bar{y}} + \left(\frac{\beta_2}{\beta_1}\right)^2 m_{\bar{x}\bar{x}} - 2\left(\frac{\beta_2}{\beta_1}\right)m_{\bar{x}\bar{y}}}{m_{vv} + \left(\frac{\beta_2}{\beta_1}\right)^2 m_{uu} - 2\left(\frac{\beta_2}{\beta_1}\right)m_{uv}} \tag{14}$$

will finally be F—distributed with $(k - 1)$ and $(N - k)$ degrees of freedom.

The ratio $\lambda = \beta_2/\beta_1$ is unknown in this expression, but if the hypothesis (6 a − c) is to be accepted it must be possible to find a real value

for λ which is compatible with an acceptable value of F, for example below the 5% limit of significance. This requirement can be expressed as follows for the variables x and y:

$$(m_{\bar{y}\bar{y}} - F_{05}m_{vv}) - 2\lambda(m_{\bar{x}\bar{y}} - F_{05}m_{uv}) + \lambda^2(m_{\bar{x}\bar{x}} - F_{05}m_{uu}) < 0 \qquad (15)$$

Putting

$$\begin{aligned} m_{\bar{y}\bar{y}} - Fm_{vv} &= c_0 \\ m_{\bar{x}\bar{y}} - Fm_{uv} &= c_1 \\ m_{\bar{x}\bar{x}} - Fm_{uu} &= c_2 \end{aligned} \qquad (16)$$

it is seen that (15) has a solution if $c_1^2 - c_0c_2 > 0$.

The finding that (15) cannot be satisfied by any real value of λ can be explained in several ways:

(1) The hypothesis (6 a − c) is correct, but the observed sample shows an excessive random deviation from the population.
(2) The variables u, v and w cannot with sufficient approximation be regarded as normally distributed.
(3) The functional relationship between the variables deviates too much from a linear relationship.
(4) There is no functional relationship between the variables, i.e. they cannot all be defined by any single family value.

3. RESULTS

In all, 84 variance ratios were obtained by analysing each school separately for family differences in the (logarithmically transformed sizes of) vaccination lesions. As could be expected with a small number of families in many of the schools, sampling errors produced a wide variation in the results. Table 1 gives the distribution of the 84 ratios according to the probability of their appearance by random chances, without influence of family factors, together with the distribution that should be expected under such conditions. A χ^2-test shows a highly significant difference between the two distributions ($P < 0.0005$) and the discrepancy originates from a predominance of large ratios in the observed distribution. It must be assumed, therefore, that the family membership has an effect on the sizes of vaccination lesions.

The corresponding tables for the variables x and y (sizes of Mantoux reactions after 10 weeks and one year) have been given in the preceding papers, and the discrepancies between observed and expected distributions were found equally significant.

TABLE 1. ANALYSIS OF LOGARITHMICALLY TRANSFORMED SIZES OF VACCINATION LESIONS FOR FAMILY DIFFERENCES

Variance ratios for 84 schools distributed according to corresponding probability fractiles (on the assumption of no family variations).

Probability fractiles for observed values of the variance ratios (percent)	Number of observed ratios in each interval	Expected number of ratios in each interval (on the assumption of no family variations)
0–10	24	8.4
10–30	17	16.8
30–50	19	16.8
50–70	11	16.8
70–90	8	16.8
90–100	5	8.4
Total	84	84.0

The next step in the analysis will be to estimate variances and covariances of all three family variables ξ, η and ζ in order to obtain quantitative expressions for the degree of their variation and to determine their interrelation in the present population of families. For this purpose, the estimated variances and covariances within and between families (m_{uu} , m_{uv} , . . . , $m_{\hat{x}\hat{x}}$, $m_{\hat{x}\hat{y}}$ . . . etc.) have been weighted by their degrees of freedom and the average values over all schools established. The results are presented in Table 2a for the variables (x, y) and Table 2b for the variables (x, z). Estimates of variances and covariances of the family values have then been computed from formula (5) and entered in the bottom lines of each table, (estimates of $\mu_{\xi\xi}$, $\mu_{\xi\eta}$ and $\mu_{\eta\eta}$ in Table 2a, of $\mu_{\xi\xi}$, $\mu_{\xi\zeta}$ and $\mu_{\zeta\zeta}$ in Table 2b). The averages given in the first line of each table (m_{uu} , m_{vv} . . . etc.) provide estimates of variances and covariances of the individual deviations from the family values (μ_{uu} , μ_{vv} , . . . etc.).

It appears that the variances of the family values roughly amount to 20-25% of the variances of the individual deviations for all three measures of response. However, as experimental errors in vaccination and in reading of reactions contribute considerably to the latter variances this ratio underestimates the importance of biological variation between families relative to the biological variation within families. An elimination of experimental errors would reduce the variance within families (μ_{uu} , μ_{vv} and μ_{ww}) but not the variance of the family values ($\mu_{\xi\xi}$, $\mu_{\eta\eta}$

TABLE 2. ESTIMATES OF VARIANCES AND COVARIANCES OF FAMILY VALUES, COMPUTED FROM ESTIMATES OF VARIANCES AND COVARIANCES WITHIN AND BETWEEN FAMILIES (AVERAGE VALUES FROM ALL SCHOOLS)

(a) Sizes of Mantoux reactions at 10 weeks and one year.

Source of variation	Degrees of freedom	Estimates of Variance		Estimates of covariance
		Mantoux reactions at 10 weeks	Mantoux reactions at one year	
Within families	497	8.99	8.37	2.77
Between families	329	13.78	13.25	6.69
Family value		2.14	2.18	1.76

(b) Sizes of Mantoux reactions at 10 weeks and logarithmically transformed sizes of vaccination lesions.

Source of variation	Degrees of freedom	Estimates of variance		Estimates of covariance
		Mantoux reactions at 10 weeks	Vaccination lesions (Log. transformed)	
Within families	1002	8.89	0.013	0.056
Between families	647	13.38	0.020	0.121
Family value		1.89	0.003	0.027

and μ_{rt}). These points have been discussed in detail in the preceding papers.

The covariances are significantly greater than zero for both pairs of family values (ξ, η) as well as (ξ, ζ). A positive correlation must, therefore, be assumed to exist both between the family values affecting post-vaccination allergy after two different intervals, as well as between the family factors affecting post-vaccination allergy and local vaccination lesions. The correlation coefficients, computed from the variances and covariances are 0.81 and 0.37, respectively.

We shall finally consider the possibility of an exact functional relation between the three family values. The normal distribution of the variables suggests that if there is such a relation, it should be approximately linear, this simplification may at least be permissible over a short interval. The hypothesis of a linear relationship is expressed

analytically in $(6\ a\ -\ c)$ and can be tested by the test indicated in the relations (14-16). The results of the tests shown in Table 3 are com-

TABLE 3. RESULTS OF TESTING A LINEAR RELATION BETWEEN THE FAMILY VARIABLES. (FOR NOTATIONS SEE EQUATIONS (14–16))

Significance limit for F	$x - y$	$x - z$	
	5%	5%	1%
c_0	3.37	0.00533	0.00459
c_1	3.54	0.0575	0.0543
c_2	3.17	3.39	2.88
$c_1^2 - c_0c_2$	+1.85	−0.0148	−0.0103

patible with a linear relation between ξ and η. It may then be tested whether the particular value $\lambda = 1$ can be accepted for these variables, it yields a variance ratio $F = 1.155$ falling in the probability interval $0.05 < P < 0.10$. The data collected in this study are thus (apart from an uninteresting constant) consistent with the hypothesis $\xi = \eta$, i.e., identity between the family values affecting allergy 10 weeks and one year after vaccination.

For the variables ξ and ζ, on the other hand, the hypothesis of linear relationship has to be rejected: we find $c_1^2 - c_0c_2 < 0$ even if we use the 1% limit of F. We must therefore reckon with the possibility that the sizes of vaccination lesions and the post-vaccination allergy depend on different (although positively correlated) family properties.

4. SUMMARY AND DISCUSSION

An analysis of variance has shown an influence of family variables on three measurable effects of BCG vaccination,—the sizes of local vaccination lesions after 10 weeks, and the level of tuberculin allergy (sensitivity) after 10 weeks, and after one year. The contribution of the family variables to the total variation of the three measures of response in the population was quite important. For each effect it was found that the variance of the family variable probably had the same order of magnitude as the variance of biological variables operating within families.

A special analysis was carried out to determine the degree of association between the three family variables defined by the different measures of responses. No significant dissociation could be demonstrated between

the two variables appearing in the allergy recorded after 10 weeks and after one year. As far as the present material shows, these two family variables can be regarded as identical or, in other words, allergy recorded after two different intervals can be assumed to depend on the *same* family property. This result is in accordance with what should be expected from a biological point of view.

In contrast to this, it was found that the sizes of vaccination lesions and the level of allergy (both recorded 10 weeks after vaccination) probably are dependent on two *different* family variables. These variables are positively correlated in the present population, but they cannot be regarded as identical. An attempt to relate this result to the common concepts of the histogenesis of the two types of reactions may be of interest.

The cellular response to tuberculin in allergic subjects is very similar to the response to tubercle bacilli, consisting mainly of a mono-nuclear cell infiltration which eventually assumes an epitheloid appearance. In fact, the tuberculin reaction is often regarded as a particular type of the Koch phenomenon in which a bacillary extract rather than bacilli is used to provoke a reaction. Again, essentially the same histological changes that occur rapidly (within 48 hours) in the Koch phenomenon can be observed after 2-3 weeks at the site of a primary tuberculous infection. It is reasonable to interpret this delayed local response in subjects without previous contact with tubercle bacilli at least partially as an allergic reaction between cells which eventually have been sensitized and tubercle bacilli still remaining at the place of injection.

According to these concepts, a positive correlation between family variables influencing sizes of vaccination lesions and post-vaccination allergy at 10 weeks was to be expected. They should both express a capacity of particular cells of the host to become sensitized to products of tubercle bacilli. It may be more surprising that, instead of a perfect functional relationship, only a slight positive correlation is found between these constitutional variables.

An important source of dissociation exists, however, in the different places of the organism from which the two effects originate. The reaction at the site of vaccination will depend on a *local* action of the bacilli and may be due partly to a sensitization of fixed histiocytic cells around the focus. The general sensitivity reflected in tuberculin reactions (and in the rapid response to reinfections) must originate from a primary stimulation of *central* organs, probably those belonging to the reticulo-endothelial system, and capable of pouring sensitized cells into the circulation for distribution to any place in the organism. The function of this system, its capacity to become sensitized and respond to

antigens in remote places does not necessarily parallel the susceptibility of the local tissue cells for sensitization. Moreover, large primary local lesions may not always be followed by a rapid dissemination of bacilli (or bacillary products), which provide the antigenic stimulus for a general sensitization. Large local lesions may even serve the purpose of localizing the bacilli and thereby preventing their spread. The defective correlation between the family variables influencing vaccination lesions and tuberculin reactions may, therefore, be related to some anatomical and pathological factors which cause a varying predominance of local and general sensitization.

The dissociation which may result from an operation of such factors can be illustrated by certain variations in the BCG vaccine, as shown in a previous study (3). Vaccines composed of dead bacilli produce little allergy, but relatively large local lesions. A high proportion of living bacilli on the other hand, favors the development of allergy. The dissociation in these cases is most naturally ascribed to a tendency of living bacilli to disseminate and of dead bacilli (and their products) to become localized at the portal of entry.

REFERENCES

(1) Palmer, Carroll E. & Nissen Meyer, Sven: *Public Health Reports, 66*: 9, 259–276, March 2, 1951.

(2) Nissen Meyer, Sven & Jensen, Chr. Munch: *Am. J. Human Genetics, 3*: 4, 325–331, December 1951.

(3) Nissen Meyer, Sven & Palmer, Carroll E.: *Bull. World Hlth. Org. 7*: 201–229, 1952.

ESTIMATION OF RELATIVE POTENCY FROM MULTIPLE RESPONSE DATA

C. Radhakrishna Rao

University of Illinois and Indian Statistical Institute

1. INTRODUCTION

When response is measured by a single variable the essential steps in the statistical treatment of data are the following (for references on this subject see Finney, 1952, bibliography),

(i) Test for parallelism of the dosage response curves to ensure the validity as dilution assay,

(ii) Test for linearity of regression to judge the appropriateness of the linear dosage response relation leading to a simple formula for the estimation of relative potency,

(iii) Test for the significance of the common regression coefficient to ensure the existence of a dosage response relation and,

(iv) The application of Fieller's theorem* in the derivation of fiducial limits of the relative potency.

The problem becomes slightly complicated when the response is measured by more than one variable. The first step is to carry out the tests (i), (ii), (iii) simultaneously for the multiple variables; this can be done by using the existing multivariate statistical tests (see references for Fisher, Hotelling, Wilks, Bartlett, and Rao in Rao, 1952, p. 271). The second step consists of the following:

(iva) Test whether an additional response measurement provides further information for the estimation of relative potency when some given measurements are already considered. This is important, because from the point of view of economy it may not be worthwhile observing a number of response measurements in addition to a few important ones, (see Rao, 1952, p. 252)

(ivb) Test whether the estimates of the relative potency from different individual response measurements are the same which is essential for a proper interpretation and estimation of relative potency,

*One of the referees of this paper comments that essentially the same method was used earlier by C. I. Bliss.

(ivc) The derivation of fiducial limits for the common value of the relative potency when (iva) and (ivb) are satisfied.

Finney (1952) gives an example of two response measurements and provides an approximate method of obtaining the fiducial limits to relative potency with the reservation that "improvement in this unsatisfactory type of statement must await further development of the statistical theory".

In this paper while illustrating tests (i), (ii), (iii) for the multivariate situation, an attempt is made to answer problems (iva), (ivb) and (ivc) in a suitable way. The problem (iva) including the adequacy of a given linear function of responses is answered by an exact test and valid fiducial limits (ivc) are obtained. The treatment of (ivb) still remains approximate and is exact only in large samples. The method of determinating fiducial limits appropriate for large samples is also discussed.

2. PRELIMINARY ANALYSIS

The following example taken out from Finney's book (Finney, 1952) refers to artificial data for an assay giving two response variates. This example is chosen because the author is not aware of any real research data on two or more response measurements and it was felt that working out a numerical example is a good way of presenting the statistical techniques.

TABLE 1. ARTIFICIAL DATA ON TWO RESPONSE MEASUREMENTS (y_1, y_2)

Dose of the standard preparation (i.u.)						Dose of the test preparation (m.g.)					
1.25		2.50		5.00		0.125		0.250		0.500	
y_1	y_2	y_1	y_2	y_1	y_2	y_1	y_2	y_1	y_2	y_1	y_2
38	51	53	49	85	47	28	53	48	48	60	43
39	55	102	53	144	51	65	53	47	51	130	50
48	46	81	46	54	39	35	52	54	48	83	48
62	51	75	51	85	41	36	54	74	50	60	51
187	203	311	199	368	178	164	212	223	197	333	192

$$\sum_{(12)} y_1 = 866, \quad \sum_{(12)} y_2 = 580 \qquad \sum_{(12)} y_1 = 720, \quad \sum_{(12)} y_2 = 601$$

$$\sum_{(24)} y_1 = 1586, \quad \sum_{(24)} y_2 = 1181$$

The first step is to obtain the analysis of dispersion (i.e. variances and covariances, Rao, 1952, p. 263) between and within doses. The formulae for between elements are

$$S_{y_1y_1} = \frac{187^2}{4} + \frac{311^2}{4} + \cdots + \frac{333^2}{4} - \frac{1586^2}{24} = 8848.84$$

$$S_{y_1y_2} = \frac{187 \times 203}{4} + \cdots + \frac{333 \times 192}{4} - \frac{1586 \times 1181}{24} = -1047.16$$

$$S_{y_2y_2} = \frac{203^2}{4} + \cdots + \frac{192^2}{4} - \frac{1181^2}{24} = 162.71$$

From the total corrected squares and products, the between elements are subtracted to obtain the within (error) elements. The second step consists of the following computations (regression analysis, sum of squares and products due to, and derivation from regression) arranged in tabular form in Tables 2.1 and 2.2 where the values of x are reduced to $-1, 0, 1$ and $-1, 0\ 1$ for both the preparations, because of the special values of x chosen in the experiment.

TABLE 2.1. SUM OF SQUARES AND PRODUCTS WITH x

Due to	S_{xy_1} (1)	S_{xy_2} (2)	S_{xx} (3)	Regression	
				(4) = (1)/(3)	(5) = (2)/(3)
(a) Standard	181	−25	8	22.625	−3.125
(b) Test	169	−20	8	21.125	−2.500
(c) Total	350	−45	16	21.875	−2.8125

TABLE 2.2. SUM OF SQUARES AND PRODUCTS FOR REGRESSION

	$S_{y_1y_1}$ (6) = (1) × (4)	$S_{y_1y_2}$ (7) = (1) × (5) = (2) × (4)	$S_{y_2y_2}$ (8) = (2) × (5)
(*a*)	4095.125	−565.625	78.125
(*b*)	3570.125	−422.500	50.000
(*c*)	7656.250	−984.375	126.562
(*d*) = (*a*) + (*b*) − (*c*)	9.000	−3.750	1.562

(d) = differences in regression (parallelism)

The entire analysis of dispersion is given in Table 3 wherein the elements due to preparations are calculated by the formulae given below. They are quite general for any number of response measurements.

$$S_{y_1y_1} = \frac{866^2}{12} + \frac{720^2}{12} - \frac{1586^2}{24} = 888.170$$

$$S_{y_1y_2} = \frac{866 \times 580}{12} + \frac{720 \times 601}{12} - \frac{1586 \times 1181}{24} = -127.750$$

$$S_{y_2y_2} = \frac{580^2}{12} + \frac{601^2}{12} - \frac{1181^2}{24} = 18.376$$

TABLE 3. ANALYSIS OF DISPERSION

Due to	D.F.	$S_{y_1y_1}$	$S_{y_1y_2}$	$S_{y_2y_2}$		
Preparations	1	888.170	−127.750	18.376		
Regression (common)	1	7656.250	−984.375	126.562		
Parallelism	1	9.000	−3.750	1.563		
Deviation from linearity	2	295.420	68.715	16.209		
Between	5	8848.84	−1047.16	162.710		
Within (error)	18	10381	749.75	205.250		
Total	23	19229.84	−297.41	367.960	Δ	Ratio to error Δ_e
		(1)	(2)	(3)	(1) × (3) − (2)²	
Error + Regression	19	18037.25	234.625	331.812	5929930	3.7805
Error + Parallelism	19	10390	746	206.813	1592270	1.0151
Error + Dev. Linearity	20	10676.42	818.465	221.459	1694500	1.0803
Error	18	10381	749.75	205.25	1568570	

All the tests considered here are based on the computations set out in Table 3. To test any component of the table with *one degree of freedom* the variance ratio is

$$F = \frac{n - p}{p}\left(\frac{\Delta}{\Delta_e} - 1\right)$$

with p and $(n - p)$ degrees of freedom where

p = number of response measurements
n = total degrees of freedom for error + the component to be tested
Δ = the determinant of the dispersion matrix of error + component to be tested
Δ_e = the above determinant for error only

For any component with *two degrees of freedom* the variance ratio is

$$F = \frac{n - p - 1}{p}\left(\sqrt{\frac{\Delta}{\Delta_e}} - 1\right)$$

with $2p$ and $2(n - p - 1)$ degrees of freedom. These two statistics are employed in the following tests.

2.1 *Test for parallelism*

$$F = \frac{19 - 2}{2}\left(\frac{\Delta}{\Delta_e} - 1\right) = \frac{17}{2}(1.0151 - 1) = 0.1283$$

This ratio is very small for 2 and 17 degrees of freedom.

2.2 *Test for deviation from linearity*

$$F = \frac{20 - 2 - 1}{2}\left(\frac{\sqrt{\Delta}}{\sqrt{\Delta_e}} - 1\right) = \frac{17}{2}(0.0040) = 0.0340$$

This value is not significant for 4 and 34 degrees of freedom.

2.3 *Test for regression*

$$F = \frac{19 - 2}{2}(2.7805) = 23.6362$$

As a variance ratio with 2 and 17 degrees of freedom the observed value is significant throwing out the possibility that one or both the regressions of y_1 on x and y_2 on x are different from zero.

2.4 *Test for additional information*

The two response measurements y_1 and y_2 may be such that one is the direct effect of the dose and the other (y_2) is a supplementary effect brought out by the first response. If this is so, the partial regression of y_2 on x when y_1 is eliminated should be zero. The value of Δ/Δ_e based on y_1 only is

$$\frac{18037.25}{10381} = 1.7375.$$

The corresponding variance ratio

$$\frac{18}{1}(0.7375) = 13.230$$

on 1 and 18 degrees of freedom is significant, showing that the regression of y_1 on x is different from zero. Consider the ratio of two values of (Δ/Δ_e) obtained for (y_1y_2) and (y_1) separately

$$3.7805 \div 1.7375 = 2.1758$$

The variance ratio for testing the significance of the partial regression is

$$\frac{17}{1}(2.1758 - 1) = 19.9886$$

with 1 and 17 degrees of freedom. This is significant showing that the second response measurement gives additional information for the estimation of relative potency. We may test the alternative hypothesis whether the response measurement y_1 is useful in addition to y_2. The value of Δ/Δ_e for y_2 alone is 1.6166. The ratio for y_1 given y_2 is $3.7805 \div 1.6166 = 2.3385$ which is significant, the corresponding variance ratio with 1 and 17 degrees of freedom being 22.7545. These tests demonstrate that an improved estimate of relative potency can be obtained by considering both the measurements instead of any one.

For other applications of such tests see Rao, 1952.

2.5 *Test for the adequacy of an assigned linear function*

We may now enquire whether a given linear function of the responses summarises the necessary information in the sense that no other linear function has non-zero regression with the dose levels independently of the given function. This means, any other response independent of the given function is not influenced by the quantity of the drug administered and will not, therefore, throw any additional information for the estimation of relative potency. The adequacy of a given function of the responses can be tested as follows.

Let $y = a_1y_1 + a_2y_2$ be the given linear function. Then the regression of y on x is computed with the help of the entries for total in Table 2.1.

$$\begin{aligned} S_{yx} &= a_1S_{xy_1} + a_2S_{xy_2} \\ &= a_1(350) + a_2(-45) \\ &= 305 \text{ for the special case } a_1 = a_2 = 1 \\ S_{xx} &= 16 \end{aligned}$$

The regression coefficient is $305/16 = 19.0625$. The sum of squares due to any category for the linear function $y = a_1y_1 + a_2y_2$ is calculated by the formula

$$a_1^2 S_{y_1y_1} + 2a_1a_2 S_{y_1y_2} + a_2^2 S_{y_2y_2}$$

where Sy_iy_j are the entries in Table 3. Thus the sum of squares due to common regression is

$$7656.250 - 2(984.375) + 126.562 = 5814.062$$

for the special case $a_1 = a_2 = 1$. Similarly the sum of squares due to error is

$$10381 - 2(749.75) + 205.25 = 9086.75$$

The ratio (Error + Regression)/(Error) for y is

$$1 + \frac{5814.062}{9086.75} = 1.6398$$

The variance ratio with 1 and 18 degrees of freedom $18(0.6398) = 11.516$ is significant. The value of Δ/Δ_e for (y_1, y_2) jointly is 3.7805 which is $3.7805 \div 1.6398 = 2.305$ times the corresponding value for y alone. The variance ratio with 1 and 17 degrees of freedom for testing its significance is $17(1.305) = 22.185$. This is significant so that the linear function $(y_1 + y_2)$ of the responses does not provide complete information on the dosage response relation.*

2.6 *Determination of the best linear function*

This leads us to the problem of determining the best linear function $(a_1y_1 + a_2y_2)$ of the responses.† The partial regression of y_1 on x when $(a_1y_1 + a_2y_2)$ is eliminated is

$$\beta_1 - k(a_1\omega_{11} + a_2\omega_{12}) \qquad (2.1)$$

where β_i is the regression of y_i on x, W_{ij} is the residual covariance of y_i and y_j and

$$k = (a_1\beta_1 + a_2\beta_2) \div \sum\sum a_ia_j\omega_{ij}$$

Equating the expression (2.1) to zero

$$\beta_1 = ka_1\omega_{11} + ka_2\omega_{12}$$

*It may be noted that in all the above tests we used the error elements based on 18 degrees of freedom from within the dose classes. Since parallelism and deviation from linearity are not significant, pooled estimates of the error elements could be obtained to have $18 + 1 + 2 = 21$ degrees of freedom. In problems of the above nature we are on the safe side in using the error based on 18 degrees of freedom.

†A similar determination seems to have been made earlier by Barnard (1935) in a problem of studying secular changes in skull characters.

Similarly calculating the partial regression for y_2 the second equation is

$$\beta_2 = ka_1\omega_{12} + ka_2\omega_{22}$$

Solving these two equations (substituting an arbitrary value for k) we obtain the ratio $a_1 : a_2$ specifying the best linear function in the sense that no other linear function of the responses has non zero partial regression with the dose. For the population parameters in the equations we can substitute their estimates and solve for a_1 and a_2 . The estimates for β_i are obtained from Table 2.1 and for ω_{ij} from the error line in Table 3.

$$21.875 = 10381a_1 + 749.75a_2$$

$$-2.8125 = 749.75a_1 + 205.25a_2$$

$$a_1 = 0.0042066, \qquad a_2 = -0.029069$$

Multiplying the coefficients by 100 (arbitrarily) the best linear function of the responses could be written

$$0.42066y_1 - 2.9069y_2$$

3. VALID FIDUCIAL LIMITS TO RELATIVE POTENCY

The preliminary tests of section 2 prepare the ground for the consideration of problems (ivb) and (ivc). We shall first take up (ivc), the problem of determining fiducial limits to λ, the relative potency and then deduce an approximate test for (ivb).

Adopting standard notation, using suffixes S and T for the constants of the standard and test preparations, consider the statistics

$$T_1 = (\bar{y}_{1T} - \bar{y}_{1S} - \lambda b_1),$$

$$T_2 = (\bar{y}_{2T} - \bar{y}_{2S} - \lambda b_2)$$

where b_1 and b_2 are the regression coefficients of y_1 and y_2 on x as obtained from the row (c) in Table 2.1.

The expectations of T_1 and T_2 are zero and the elements

$$\frac{T_1^2}{\mu}, \frac{T_1T_2}{\mu}, \frac{T_2^2}{\mu}$$

where

$$\mu = \frac{1}{n_1} + \frac{1}{n_2} + \frac{\lambda^2}{S_{xx}}$$

n_1 = sample size for the test preparation
n_2 = sample size for the second preparation
S_{xx} = the entry in column (3) for total in Table 2.1

estimate the same quantities as the error elements. If the error elements are denoted by

$$W_{11}\ ,\ W_{12}\ ,\ W_{22}$$

with degrees of freedom k, then the statistic

$$\frac{\begin{vmatrix} W_{11} + \dfrac{T_1^2}{\mu} & W_{12} + \dfrac{T_1T_2}{\mu} \\ W_{12} + \dfrac{T_1T_2}{\mu} & W_{22} + \dfrac{T_2^2}{\mu} \end{vmatrix}}{\begin{vmatrix} W_{11} & W_{12} \\ W_{12} & W_{22} \end{vmatrix}} - 1 \tag{3.1}$$

multiplied by $(k - 1)/2$ is a variance ratio with 2 and $(k - 1)$ degrees of freedom. Equating the above to $2(5\%$ value of $F)/(k - 1)$ we obtain a quadratic in λ giving two roots. These are valid fiducial limits to λ. The equation can be written

$$W_{22}T_1^2 - 2W_{12}T_1T_2 + W_{11}T_2^2 = \frac{2\mu}{k - 1} F_{5\%}\Delta_e \tag{3.2}$$

In our example

$$T_1 = -12.1667 - 21.875\lambda, \qquad T_2 = 1.7500 + 2.8125\lambda$$

$$\mu = \frac{1}{12} + \frac{1}{12} + \frac{\lambda^2}{16}$$

and W_{ij} are the error elements in Table 3. The equation is

$$27258\lambda^2 + 320155\lambda + 94101.6 = (11534\lambda^2 + 30756)(F_{5\%} = 3.5914)$$

$$= 41423\lambda^2 + 110457 \tag{3.3}$$

or

$$231162\lambda^2 + 320155\lambda - 16355.4 = 0$$

giving two roots

$$\lambda_1 = -1.4343, \qquad \lambda_2 = 0.04934$$

The fiducial limits for relative potency are

$$R_1 = 10 \text{ antilog } (\lambda_1 \log_{10} 2) = 10 \text{ antilog } (\bar{1}.5682)$$

$$= 3.700 \tag{3.4}$$

$$R = 10 \text{ antilog } (.01485) = 10.348$$

If the equation (3.3) has only imaginary roots, then there is an indication that the relative potencies as determined by the two responses are different and the question of fiducial limits to common relative potency does not arise.

We may now compare these limits with the limits obtained by using the first measurement alone. The quadratic to be solved is

$$(\bar{y}_{1S} - \bar{y}_{1T} - \lambda b_1)^2 = \left(\frac{2}{n} + \frac{\lambda^2}{S_{xx}}\right)\frac{W_{11}}{18} F_{5\%}$$

or inserting the numerical values ($F_{5\%}$ having 1 and 18 d.f)

$$319.4160\lambda^2 + 532.2930\lambda - 276.2371 = 0$$

which has the two roots

$$\lambda_1 = -2.0818, \qquad \lambda_2 = 0.4154$$

The fiducial limits

$$R_1 = 10 \text{ antilog } (\bar{1}.37332) = 2.346$$

$$R_2 = 10 \text{ antilog } (0.12504) = 13.337$$

are much wider than the limits based on both the response measurements.

4. TEST FOR THE EQUALITY OF RELATIVE POTENCIES

The estimated limits of section 3 cease to have a meaning if the relative potencies relevant to the two response measurements differ. In fact, such a difference would make the fiducial limits wider and this is an indication that our assumption of equality is not valid. An objective test of this hypothesis would be necessary to justify the computations of section 3. No exact test could be found but the following test appears to be good enough for practical application. If difference is detected, then the validity of the assay is open to question.

The statistic considered in (3.1)

$$(W_{22}T_1^2 - 2W_{12}T_1T_2 + W_{11}T_2^2) \div \mu\Delta_e \qquad (4.1)$$

was used in constructing the variance ratio with 2 and 17 degrees of freedom. We may find the value of λ for which (4.1) is a minimum. This value provides an intutively good point estimate of the common relative potency. Substituting the numerical values of section 3 the expression to be minimised is (using the computations of 3.3)

$$\frac{272585\lambda^2 + 320155\lambda + 94101.6}{98039\lambda^2 + 261426} = \frac{p\lambda^2 + q\lambda + r}{g\lambda^2 + h}$$

The equation giving the stationary values of λ is

$$-qg\lambda^2 + 2(ph - gr)\lambda + gh = 0$$

or

$$\lambda^2 - 2\left(\frac{p}{q}\frac{h}{g} - \frac{r}{q}\right)\lambda - \frac{h}{g} = 0$$

$$\lambda^2 - 2(1.97642)\lambda - 2.666551 = 0$$

One of the roots is

$$-0.5873^* \qquad \text{(with the point estimate 6.656)}$$

leading to the minimum value of the variance ratio

$$\frac{2p\lambda + q}{2g\lambda} = \frac{p}{g} + \frac{q}{g}\frac{1}{2\lambda}$$

$$= 2.7804 - \frac{3.2656}{1.1746}$$

$$= 2.7804 - 2.7802 = .0002$$

In large samples 18 times this quantity is distributed as χ^2 with 1 degree of freedom but in small samples 18 times this quantity can be considered as a variance ratio with 1 and 18 degrees of freedom. The computed value 0.0036 is incredibly small showing that the two estimates agree remarkably well; the artificial data seem to have been constructed with some ingenuity!

The analysis is presented here in such a way that generalisation to more than 2 response measurements is automatic.

5. LARGE SAMPLE FIDUCIAL LIMITS TO RELATIVE POTENCY

5.1 *Exact limits when the regression coefficients are known*

Let us consider the special case when β_1 and β_2 the two regression coefficients are known. The two statistics

$$t_1 = (\bar{y}_{1T} - \bar{y}_{1S} - \lambda\beta_1), \qquad t_2 = (\bar{y}_{2T} - \bar{y}_{2S} - \lambda\beta_2)$$

have zero expectation and the determinantal ratio corresponding to them is

$$\frac{1}{\Lambda} = \begin{vmatrix} W_{11} + \frac{t_1^2}{\nu} & W_{12} + \frac{t_1 t_2}{\nu} \\ W_{12} + \frac{t_1 t_2}{\nu} & W_{22} + \frac{t_2^2}{\nu} \end{vmatrix} \div \begin{vmatrix} W_{11} & W_{12} \\ W_{12} & W_{22} \end{vmatrix} \qquad (5.1)$$

*The point estimate, as the ratio of difference in means to the regression coefficient for the best linear function determined earlier, agrees with the above to four decimal places.

where $v = (1/n_1 + 1/n_2)$. The statistic

$$M = \beta_2(\bar{y}_{1T} - \bar{y}_{1S}) - \beta_1(\bar{y}_{2T} - \bar{y}_{2S})$$

supplies ancillary information on λ and the variance ratio with 1 and 18* degrees of freedom for testing that M has zero expectation (which implies a test for the equality of relative potencies for the two responses) is $18(\Lambda_1^{-1} - 1)$ where Λ_1^{-1} is

$$\Lambda_1^{-1} = 1 + \frac{M^2}{\nu(\beta_2^2 W_{11} - 2\beta_1\beta_2 W_{12} + \beta_2^2 W_{22})} \tag{5.2}$$

The test of the hypothesis for any specified λ is supplied by the variance ratio

$$\frac{17}{1}\left(\frac{\Lambda^{-1}}{\Lambda_1^{-1}} - 1\right) \tag{5.3}$$

which has 1 and 17 degrees of freedom. Equating this to the 5% value of F we obtain a quadratic in λ giving the exact fiducial limits. This method of determining the fiducial limits is quite general and is applicable to cases where a number of p correlated normal estimates of the same parameter are available giving rise to $(p - 1)$ ancillary statistics in the form of differences. We can find the fiducial limits to the parameter by considering the conditional distribution of any other statistic given the ancillaries. A typical example is that of determining the common mean of p correlated normal variables $(x_1 \cdots , x_p)$ on the basis of a sample of size n from a p-variate population. This is equivalent to determining the fiducial limits to the parameter α in the regression equation

$$x_p = \alpha + \beta_1 y_1 + \cdots + \beta_{p-1} y_{p-1}$$

where

$$y_i = x_i - x_p , \qquad i = 1, \cdots , p - 1$$

are considered fixed.

5.2 *Fiducial Limits in Large Samples*

In the present problem the above method cannot be used as the regression coefficients are unknown. The following analogous procedure is useful in cases where β_i are unknown and the sample size is large. The determinantal ratio in (3.1) is

$$1 + \frac{W_{22}T_1^2 - 2W_{12}T_1T_2 + W_{11}T_2^2}{\mu\Delta_e} \tag{5.4}$$

*Since β are known, improved estimates of error elements could be found so that the degrees of freedom will be more than 18 in general. This refinement is ignored here.

The minimum value of this as found in the last section is 1 + (0.0002) and this provided a test for the equality of relative potencies. The ratio of the expression (5.4) to the minimum value (1.0002) is

$$\Lambda_2^{-1} = \frac{W_{22}T_1^2 - 2W_{12}T_1T_2 + W_{11}T_2^2 + \mu\Delta_e}{(1.0002)\mu\Delta_e} \tag{5.5}$$

The statistic

$$\frac{17}{1}(\Lambda_2^{-1} - 1) \tag{5.6}$$

when the sample size is large is a variance ratio with 1 and 17 degrees of freedom. Equating (5.6) to the 5% value 4.45, the fiducial limits to λ are obtained. The equation is

$$\Lambda_2^{-1} = 1 + \frac{4.45}{17} = 1.26176$$

$$W_{22}T_1^2 - 2W_{12}T_1T_2 + W_{22}T_2^2 = \mu\Delta_e(1.2618 \times 1.0002 - 1)$$
$$= 0.2621\mu\Delta_e$$

This reduces to (using the computations already carried out)

$$246889\lambda^2 + 320155\lambda + 25581.8 = 0$$

giving the two roots

$$\lambda_1 = -1.2112, \qquad \lambda_2 = -0.0855$$

The fiducial limits are

$$R = 10 \text{ antilog } (\lambda_1 \log_{10} 2) = 10 \text{ antilog } (\bar{1}.6354)$$
$$= 4.320$$
$$R = 10 \text{ antilog } (\bar{1}.9743) = 9.425$$

These are much narrower than the valid limits obtained in (3.4). It must be remembered that the above limits are approximate and in large samples there should be good agreement between the two.

REFERENCES

1. Barnard, N. N. (1935): The secular variation of skull characters in four sereis of Egyptian skulls. *Ann. Eugen., Lond., 7*, 89.
2. Finney, D. J. (1952): *Statistical Method in Biological Assay.* Charles Griffn Company, London.
3. Rao, C. R. (1952): *Advanced Statistical Methods in Biometric Research.* John Wiley & Sons, New York.

ERROR OF THE DETERMINATION OF THE EOSINOPHIL COUNT IN PERITONEAL FLUID OF THE RAT

PETER B. DEWS, M.B., CH.B.
Research Associate,
Division of Biometry and Medical Statistics
GEORGE M. HIGGINS, PH.D.
Section of Anatomy,
Mayo Foundation, University of Minnesota
AND
JOSEPH BERKSON, M.D., D.Sc.
Division of Biometry and Medical Statistics,
Mayo Clinic, Rochester, Minnesota

The object of this investigation was to determine the error of enumeration of eosinophil cells in samples of peritoneal fluid from rats, using the conventional technic employing a dilution pipet and counting chamber.

Berkson, Magath and Hurn (1940)[1] studied the error of count in human blood and came to the conclusion that the error of the leukocyte count represented as the coefficient of variation is given by the formula

$$V = \sqrt{\frac{100^2}{n_b} + \frac{4.6^2}{n_c} + \frac{4.7^2}{n_p}} \qquad (1)$$

where V is the coefficient of variation of the count, that is, the standard deviation expressed as a percentage of the mean; n_b is the total number of cells counted; n_c is the number of hemocytometer chambers used; 4.6 per cent is the error of the chamber; n_p is the number of pipets used; 4.7 per cent is the error of the pipet.

Chamberlain and Turner (1952)[2] recently have reinvestigated the problem; they agree with the form of the formula proposed by Berkson and associates but found somewhat different constants for the chamber and pipet errors. Their formula is

$$V = \sqrt{\frac{100^2}{n_b} + \frac{3.72^2}{n_c} + \frac{7.43^2}{n_p}} \qquad (2)$$

where the symbols have the same meaning as in (1).

In the situation usually obtaining in routine practice, in which only one pipet and one chamber are used for each count, the formula of Chamberlain and Turner gives somewhat higher values for the coefficient of variation than does the formula of Berkson and associates. For this situation, the formulas of Berkson and associates and of

Chamberlain and Turner are respectively

$$V = \sqrt{\frac{100^2}{n_b} + 43.25} \qquad \text{and} \qquad V = \sqrt{\frac{100^2}{n_b} + 69.04}$$

It was of interest to determine whether these formulas, derived from studies on total leukocyte counts of human blood, were applicable to counts made of a single variety of cell (the eosinophil), in a different fluid (peritoneal fluid) of a different species (the rat) and using a different counting chamber (the Fuchs-Rosenthal instead of the Neubauer).

A sample of peritoneal fluid was taken from a normal rat by a technic to be described elsewhere (Higgins, 1952)[3] and placed on a siliconed microscope slide. Then three people each drew a sample of the fluid to the 0.1 mark of a Thoma-Ziess white cell pipet; the three samples were taken in quick succession, and the fluid was stirred with the tip of the pipet before each sampling to prevent settling of the cells*. The pipet was then filled to the appropriate mark with the phloxine-propylene glycol-water mixture recommended by Randolph (1944)[4] to give a 1:100 dilution of the peritoneal fluid, shaken at least thirty seconds, allowed to stand at least fifteen minutes, reshaken, and then, after rejection of the first three drops issuing, used to fill a Fuchs-Rosenthal counting chamber. The cells were allowed to settle and then the number of eosinophils in both sides of the chamber were counted directly. Since the volume of fluid over the rulings on each side of the chamber is 3.2 mm.^{3}, and since the peritoneal fluid was diluted 1:100, the estimated number of cells/mm.^{3} of peritoneal fluid is equal to the number of cells counted multiplied by 15.625. The same three observers made all the counts. The order in which the three counters took their samples of fluid from the drop on the slide was determined from a book of random numbers, with the provision that each counter sampled first, second, and third in order an equal number of times. Thus three parallel counts were obtained, one by each of the counters using a single pipet and chamber, on specimens of peritoneal fluid from 60 rats.

RESULTS

The over-all mean counts for the 60 rats of the samples that were taken first, second, and third were 328, 323, and 331 respectively**. There is clearly no evidence of any settling of the cells during the time period between the taking of the first and third samples; this is not

*Preliminary exploratory experimentation disclosed that there was a settling of cells with lapse of time, but that within the short time required for three successive samples to be taken with the precautions mentioned, it could be considered that the fluid sampled by the three was a uniformly mixed identical specimen.

**Except where otherwise stated, the text and tables refer to numbers of cells actually counted, and not to the estimated number per mm.^{3}

surprising in view of the stirring of the fluid and the fact that the whole operation took less than thirty seconds.

Preliminary to an estimate of the error of the count from the sixty sets of three counts each, an examination was made of the counts to ascertain whether there was any evidence of bias in the counting of the two sides of the chamber field. If O_1 represents the count made on one side and O_2 the count made on the other, then if the cells are randomly distributed over the two sides, the quantity

$$\text{Chi}^2 = \frac{(O_1 - O_2)^2}{O_1 + O_2}$$

should be distributed closely as Chi^2 for one degree of freedom. In two of the 180 counts (3×60), the counts on the two sides of the chamber were not recorded separately, leaving 178 pairs to be considered. For each of these the Chi^2 "P" was determined as for one degree of freedom*. If the distribution of the Chi^2 observed followed the Chi^2 distribution of 1 D.F., there should be an equal number of "P" values in each of the ten intervals $0 - 0.1, 0.1 - 0.2, \ldots, 0.9 - 1.0$. Berkson and associates (1935)[5] and Lancaster (1950)[6] have used this distribution to test whether the counts were in reasonable agreement with unbiasedness of the counting in the individual chambers. The distribution of the P's is shown in table 1.

It will be noticed that there is some excess of values of P greater than 0.5, indicating that the counts from the two sides of the hemocytometer chamber agreed more closely, on the average, than would have been expected. However, the deviation from expectation is not great and the total $\chi^2 = 7.06$ for the distribution of the "P's" is not significantly smaller than its expectation of 9, corresponding to nine degrees of freedom.

Each of the sets of three counts made from the peritoneal fluid of an individual rat furnished an estimate based on two degrees of freedom of the standard deviation of the count

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{2}}.$$

The standard deviation so estimated, divided by the mean obtained from the three counts ($\bar{x} = \sum x/3$) was used as an observation of the coefficient of variation of the count. Also the mean of the counts was inserted into the formula of Berkson and associates as well as into the

*The "P" values can be obtained from a table of the normal curve; the "P" required is twice the area of the unit normal curve beyond the normal deviate evaluated as n.d. $= \sqrt{\chi^2}$.

TABLE 1

DISTRIBUTION OF P's

	$P_{x's_1}$				
	0–0.1	0.1–0.2	0.2–0.3	0.3–0.4	0.4–0.5
Observed	12	16	22	12	17
Expected	17.8	17.8	17.8	17.8	17.8
$\frac{(\text{Obs.} - \text{Exp.})^2}{\text{Exp.}}$	1.89	0.18	0.99	1.89	0.04

	$P_{x's_1}$					
	0.5–0.6	0.6–0.7	0.7–0.8	0.8–0.9	0.9–1	Total
Observed	19	18	18	22	22	178
Expected	17.8	17.8	17.8	17.8	17.8	178
$\frac{(\text{Obs.} - \text{Exp.})^2}{\text{Exp.}}$	0.08	0.00	0.00	0.99	0.99	7.06

formula of Chamberlain and Turner, and in this way the respective formulary estimates of the coefficient of variation were obtained. A comparison of the averages of these coefficients of variation, the observed and the formulary estimates, is shown in table 2 separately for the mean counts below and above the median as well as for the total series of observation.*

COMMENT

The results (table 2) corroborate, for the estimate of the eosinophil count, the finding first clearly demonstrated by Berkson and associates, that when the blood count is estimated with the usual type of hemocytometer and diluting pipet, the manipulations with these required to accomplish the count add considerably to the imprecision of the count arising from the Poisson variability within the hemocytometer field.

*The mean of the estimated number of eosinophils per mm.^{3} in the peritoneal fluids of the 60 rats studied here was 5,109 cells. The range was from 927 to 20,198, and the estimated standard deviation was 4,125.

TABLE 2

MEANS OF COEFFICIENTS OF VARIATION

Group	No. of rats	Mean no. of cells counted	Means of coefficients of variation of counts as estimated from			
			"Observed"	Formula of		
				"Poisson"	Berkson	Chamberlain
Mean count below median	30	164	10.1	8.2	10.5	11.7
Mean count above median	30	490	8.5	5.0	8.3	9.8
Total	60	327	9.3	6.6	9.4	10.7

There is a remarkably close agreement between the average coefficients of variation estimated from counts made in triplicate on the peritoneaf fluid of rats and the average of the values given by the formula ol Berkson and also by the formula of Chamberlain, the closest agreement being with the former.

SUMMARY

Triplicate eosinophil counts on single samples from rats, of peritoneal fluid containing eosinophils of the order of 5,000 per mm.^{3} were performed. The average coefficient of variation of the counts was 9.3 per cent, in close agreement with formulary estimates of the coefficient of variation. Since it is common practice to consider an estimate significantly determined within $\pm$ 2 S.E., this means that using a single pipet and counting chamber and a dilution of 1:100, an eosinophil count of 5,000 per mm.^{3} will be significantly determined within about $\pm$ 20 per cent. A graph (fig. 1) is given permitting the coefficient of variation as predicted by the formula of Berkson and associates to be read directly.

We wish to thank the Misses Dorothy Failor, Betty Ann Hennessey and Mary Woods for their technical assistance in carrying out these experiments.

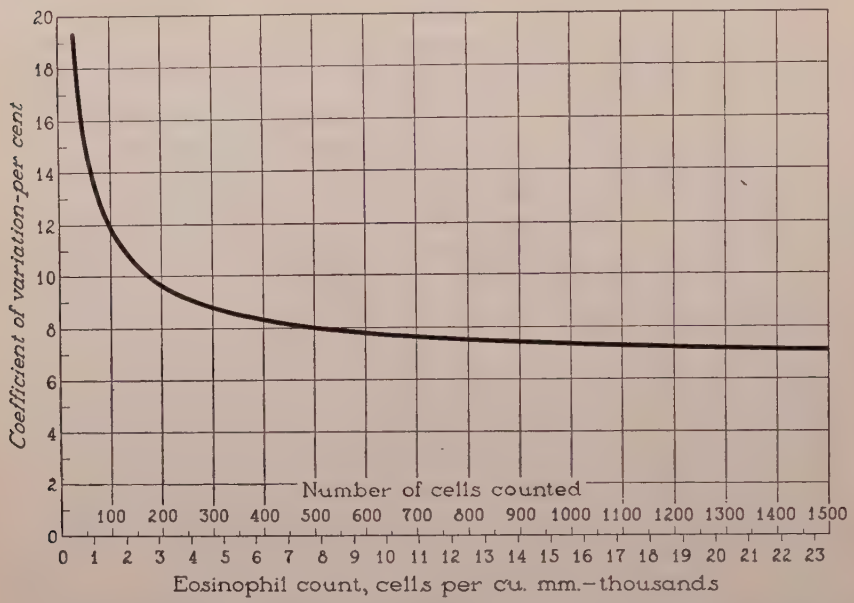

FIG. 1. The curve gives the coefficient of variation (that is, the standard deviation expressed as a percentage of the mean) evaluated by the formula of Berkson and associates, for different numbers of cells counted, using one pipet and one hemocytometer for each count. The corresponding figures for eosinophil count in cells per cubic millimeter apply when a 1:100 dilution has been used in estimating the count.

It is usual practice to consider an estimate as determined significantly within ± 2 s.e., so that in stating the error of a count, the coefficient of variation as given on the graph should be multiplied by 2 to give the percentage error.

REFERENCES

1. Berkson, Joseph, Magath, T. B., and Hurn, Margaret: The error of estimate of the blood cell count as made with the hemocytometer. *Am. J. Physiol. 128*: 309–323, 1940.
2. Chamberlain, A. C., and Turner, F. M.: Error and variations in white cell counts. *Biometrics. 8*, 55–65, 1952.
3. Higgins, G. M.: Eosinophil in peritoneal exudates. In: Proceedings of Conference on the Morphology and Physiology of the Eosinophil Leukocyte. Bar Harbor, Maine, 1952. (In press.)
4. Randolph, T. G.: Blood studies in allergy. I. The direct counting chamber determination of eosinophils by propylene glycol aqueous stains. *J. Allergy. 15*: 89–96, 1944.
5. Berkson, Joseph, Magath, T. B., and Hurn, Margaret: Laboratory standards in relation to chance fluctuations of the erythrocyte count as estimated with the hemocytometer. *J. Am. Stat. Assoc. 30*: 414–426, 1935.
6. Lancaster, H. O.: Statistical control in haematology. *J. Hyg. 48*: 402–417, 1950.

HOW MANY ORGANISMS?

JANE WORCESTER

Harvard School of Public Health

Two quite distinct statistical methods are in use at the present time for estimating without a direct count the amount of an infectious agent present in a suspension. Since the assumptions underlying the two methods are different and since neither set of assumptions may be fulfilled in practice, it seems worthwhile to contrast the two methods and to point out some of the difficulties which arise when an attempt is made to proceed under a set of assumptions which appears a priori to be more reasonable. The two methods which are to be compared are the estimation of fifty per cent end points and the estimation of densities by means of the "most probable number".

Suppose, for example, one wishes to determine the amount of an infectious agent present in a suspension from the response of suitably chosen experimental animals. Serial dilutions, generally logarithmic, are made from the original suspension. Groups of animals are inoculated with a standard amount at each dilution. After a suitable length of time, the number failing to respond is recorded and the percentage of failures is computed at each dilution. This procedure results in a series of percentages which tend to increase as the dilutions increase. It is from this series of percentages that the strength of the suspension in terms of the fifty per cent end point or the "most probable number" is estimated.

Both the integrated normal and the logistic curves have been used for the estimation of fifty per cent end points. These dosage response curves have been fitted by the methods of maximum likelihood, least squares and minimum Chi-square. While each curve (1) and each method of fitting has its ardent proponents (2) it matters very little in a given experiment which is used. Theoretically, the following assumptions should be satisfied before the constants of either curve are calculated. (A) The number of organisms inoculated into each animal at a given dilution is the same. Stated another way, this assumption means that the number of organisms inoculated into the animals must be large enough so that the error introduced by the random distribution of the organisms in the original suspension and in the samples at the various dilutions is small relative to the differences on the dilution scale. (B) The susceptibilities of the animals to the agent are dis-

tributed normally or on the derivative of the logistic. The dosage response curve arising under these assumptions has a steep slope if the animals are homogeneous with respect to susceptibility and a more gradual slope if the variation in susceptibility is large. In other words, for a given set of dilutions, the slope of the dosage response curve is determined by the distribution of susceptibilities in the animals. Chance variation arises from the number of animals at each dilution and introduces variability about the curve. The constants which are estimated are the fifty per cent point (in units on the dilution scale) which is a measure of the strength of the suspension and the slope of the dosage response curve (in probit or logit units). The latter constant may be interpreted as a measure of the susceptibility of the animals.

If the strength of the suspension is to be determined by the use of the "most probable number", (3) the experimental procedure is essentially the same. The density of the original suspension is determined under the following assumptions. (1) The organisms are distributed at random throughout the suspension and at each dilution made from it. Under this assumption samples at a particular dilution do not contain the same number of organisms. Indeed the number of organisms per sample is assumed to follow the law of small numbers.* (2) Each sample when inoculated into an animal produces response if the sample contains one or more organisms. The animals, in other words, are assumed to be homogeneous with respect to susceptibility. The single parameter which is generally estimated under this set of assumptions is the density in units of number of organisms in the original suspension or at a specified dilution. However, the fifty per cent point can be computed directly from the density, if it is wanted for comparative purposes.

The "most probable number" has been used for many years for estimating the number of organisms in water and in milk. The suitably chosen experimental animal has been culture medium in a test tube. Comparisons of the number of organisms obtained by direct count with those obtained under the "most probable number" theory have shown good agreement. There has been little, if any, evidence to suggest that

*The question may be raised as to how a given inoculum was obtained. If, for example, samples of size v are taken from V ml. in which there are b organisms and if from the samples of size v sub-samples of size d are made, the probability of failure to respond becomes

$$P = \exp\left\{\frac{-vb}{V}\left(1 - e^{\frac{-d}{v}}\right)\right\}.$$

However, if samples of size d are drawn directly from V ml.,

$$P = \exp\left\{\frac{-db}{V}\right\};$$

The latter expression is the one routinely assumed under the "most probable number" theory.

the assumptions are violated. The situations under discussion involve the necessary substitution of a living animal for a test tube. Under these conditions, where, for example, viruses or rickettsiae are being studied, direct counts of viable organisms are impossible or at best impractical. In some of these cases the individuals working with the organisms believe that the number necessary to produce response is small. If this be so, random variation in the number of organisms in samples from a given dilution becomes important relative to the dilution scale and must be taken into consideration. At the same time, they believe that the animals vary in susceptibility. Therefore, it becomes necessary to postulate variation both in the number of organisms in samples from a given dilution and in the response of different animals to the same dose. Estimation of the "most probable number" or the fifty per cent end point is accomplished by assuming one or the other of these to be constant.

It is of some interest to see if these assumptions can be incorporated into a workable theory. Let y be the number of organisms present in samples at a specified dilution of the original suspension. Let the average number of organisms present in samples taken at the d_i dilution be $d_i y$. If the organisms are distributed at random and the law of small numbers holds, the fractions of samples at the d_i dilution with 0, 1, 2, etc. organisms will be:

Number of organisms	*Fraction of samples*
0	$e^{-d_i y}$
1	$e^{-d_i y}(d_i y)$
2	$e^{-d_i y}(d_i y)^2/2!$
—	
—	
—	
x	$e^{-d_i y}(d_i y)^x/x!$

Now let it be assumed that the probability of response in an animal varies with the number of organisms it receives according to $x/(1 + x)$, where x is the number of organisms. The probability of not responding would then be $1/(1 + x)$. (This replaces the "most probable number" assumption that an animal responds if it receives at least one organism.)

The probability of response, Q_{d_i}, at dilution d_i becomes

$$Q_{d_i} = e^{-d_i y}\left[d_i y \frac{1}{1+1} + \frac{(d_i y)^2}{2!}\frac{2}{1+2} + \cdots + \frac{(d_i y)^x}{x!}\frac{x}{1+x} + \cdots\right].$$

Fortunately this series can be summed and

$$Q_{d_i} = 1 - \frac{1}{d_i y}[1 - e^{-d_i y}], \qquad P_{d_i} = \frac{1}{d_i y}[1 - e^{-d_i y}].$$

If n_i is the number of animals inoculated at dilution d_i and if s_i fail to respond and $n_i - s_i$ respond, the probability of the observed results becomes

$$\text{constant } Q_{d_1}^{n_1 - s_1} P_{d_1}^{s_1} Q_{d_2}^{n_2 - s_2} P_{d_2}^{n_2 - s_2}, \text{ etc.}$$

and the maximum likelihood estimate of y results from solving the equation

$$\sum \frac{d_i(n_i - s_i)(1 - e^{-d_i y})}{d_i y - (1 - e^{-d_i y})} + \sum \frac{d_i s_i e^{-d_i y}}{1 - e^{-d_i y}} - \sum \frac{n_i}{y} = 0.$$

The standard error of y can be obtained in the usual manner from the second derivative of the likelihood which is

$$\frac{d^2 L}{dy^2} = -\sum \frac{n_i}{y^2 P_{d_i} Q_{d_i}}[1 - e^{-d_i y} - Q_{d_i}]^2.$$

The assumption that the animals respond according to the expression $x/(1 + x)$ is, in fact, only a little less arbitrary than the assumption it replaced since it implies that all types of organisms affect all types of animals in the same way. What is needed in the equation relating the response of the animals to the number of organisms is a parameter to be estimated from the data and which is therefore peculiar to the experimental situation at hand. The number of expressions which might be tried is infinite. The probability of response could be set equal to $1 - q^x$, where, as before, x is the number of organisms and q is a constant which can be interpreted as the probability of failing to respond to one organism. If the law of small numbers is again assumed, the probability of response at dilution d_i becomes

$$Q_{d_i} = 1 - e^{-d_i y(1-q)}.$$

Unfortunately the method of maximum likelihood does not give estimates for y and q but gives only a value for $y(1 - q)$. This result is not without interest since this product, $y(1 - q)$, is identically the same as the value obtained for the density under the assumptions leading to the "most probable number". Indeed, $q = 0$ leads directly to the "most

probable number" result. However, the interpretation of this value as a product is perhaps closer to the facts.

Another equation which might be used for the probability of response is $x/(b + x)$. This expression, for positive values of b, gives mortalities between 0 and 100%. If b is between 0 and 1, the curve is higher than $x/(1 + x)$ and if b is greater than 1, it is lower. The proportion responding at dilution d_i becomes

$$Q_{d_i} = e^{-d_i y}\left[d_i y \frac{1}{b+1} + \frac{(d_i y)^2}{2!}\frac{2}{b+2} + \cdots + \frac{(d_i y)^x}{x!}\frac{x}{b+x} + \cdots\right].$$

To date this series has been summed only for integral values of b.

An expression for the probability of response which leads to a result where estimates of both parameters can be obtained is $c(x)/(1 + x)$ where c is the probability of responding to a large number of organisms and is a positive number between 0 and 1, although there is nothing in the method of fitting which ensures this. This expression and the law of small numbers make

$$Q_{d_i} = c\left[1 - \frac{1}{d_i y}(1 - e^{-d_i y})\right].$$

The method of maximum likelihood leads to two equations which must be solved for c and y.

$$\sum \frac{d_i(n_i - s_i)(1 - e^{-d_i y})}{d_i y - (1 - e^{-d_i y})} + \sum \frac{d_i s_i[1 - c(1 - e^{-d_i y})]}{d_i y - c[d_i y - (1 - e^{-d_i y})]} - \sum \frac{n_i}{y} = 0,$$

and

$$\sum \frac{n_i - s_i}{c} - \sum \frac{s_i[d_i y - (1 - e^{-d_i y})]}{d_i y - c[d_i y - (1 - e^{-d_i y})]} = 0.$$

The standard errors of the constants may be computed from the second derivatives of the likelihood.

$$\frac{\partial^2 L}{\partial c^2} = -\frac{1}{c^2}\sum \frac{n_i Q_{d_i}}{P_{d_i}},$$

$$\frac{\partial^2 L}{\partial c\,\partial y} = -\frac{1}{y}\sum \frac{n_i(1 - e^{-d_i y})}{P_{d_i}} + \frac{1}{cy}\sum \frac{n_i Q_{d_i}}{P_{d_i}},$$

$$\frac{\partial^2 L}{\partial y^2} = -\frac{1}{y^2}\sum \frac{n_i[Q_{d_i} - c(1 - e^{-d_i y})]^2}{P_{d_i} Q_{d_i}}.$$

TABLE I.

BASIC DATA

Dilution	d_i	Species A				Species B			
		n_i	s_i	$n_i - s_i$	s_i/n_i	n_i	s_i	$n_i - s_i$	s_i/n_i
$10^{-7.5}$	10	28	3	25	11%	15	1	14	7%
$10^{-8.0}$	3.162	29	6	23	21	17	1	16	6
$10^{-8.5}$	1.	36	14	22	39	16	2	14	12
$10^{-9.0}$	.3162	38	25	13	66	17	15	2	88
$10^{-9.5}$	.1	26	24	2	92	18	16	2	89

TABLE II.

DESCRIPTIVE CONSTANTS

1) The logistic (with the logarithm of the dilution as the x scale) where $P_{x_i} = \frac{1}{2} - \frac{1}{2} \tanh \alpha(x_i - \gamma)$

	Species A	Species B
50% point (γ)	$-8.637 \pm .089$	$-8.745 \pm .094$
slope (α)	$1.084 \pm .177$	$1.735 \pm .352$
$r_{\alpha\gamma}$	.062	.033

2) The "most probable number" where $P_{d_i} = e^{-d_i y}$.

	Species A	Species B
Number of organisms at $10^{-8.5}(y)$	$.542 \pm .079$	$.385 \pm .080$
50% point	$-8.392 \pm .063$	$-8.244 \pm .090$

3) The response curve, $x/(1 + x)$, where $P_{d_i} = (1/d_i y)\,[1 - e^{-d_i y}]$.

	Species A	Species B
Number of organisms at $10^{-8.5}(y)$	$2.104 \pm .379$	$2.866 \pm .730$
50% point	$-8.620 \pm .078$	$-8.755 \pm .111$

4) The response curve $c[x/(1 + x)]$ where $P_{d_i} = 1 - c[1 - (1/d_i y)(1 - e^{-d_i y})]$

	Species A	Species B
Number of organisms at $10^{-8.5}(y)$	$2.815 \pm .735$	$2.878 \pm .908$
c	$.916 \pm .058$	$.997 \pm .051$
r_{yc}	$-.626$	$-.572$
50% point	$-8.680 \pm .132$	$-8.755 \pm .149$

It is of some interest to compare these results. Following is part of an experiment designed to see if species A and species B behave in the same way with respect to a particular infectious agent (4). The basic information is in Table I. The column d_i has been referred to the $10^{-8.5}$ dilution.

The percentage of animals failing to respond appears to increase more abruptly in Species B than in Species A. However, comparisions of the percentages at the five dilutions by the χ^2-test, a dubious procedure in view of the size of some of the numbers, show none of the differences to be significant at the .05 level. The sum of the five χ^2 values gives $P = .12$. The reactions of the species to the agent are not remarkably unlike, but the question here relates not to this difference but rather to the number of organisms involved.

The various assumptions lead to the values shown in Table II. The 50% points are in the dilution scale and are exponents of 10. The standard errors of the 50% points in methods 2, 3 and 4 were estimated by differentiating the expressions for P_{d_i} , squaring and substituting the maximum likelihood estimates of the parameters, their variances and covariances.

The differences, observed s_i minus expected s_i , for the two species at the several dilutions are given in Table III.

TABLE III.

OBSERVED s_i MINUS EXPECTED s_i

Dilution	Species A				Species B			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
$10^{-7.5}$	.8	2.9	1.7	− .3	.8	.7	.5	.4
$10^{-8.0}$	.2	.8	1.6	.6	−.2	−4.0	−.9	−.9
$10^{-8.5}$	−1.3	−6.9	−1.0	− .0	−2.8	−8.9	−3.3	−3.3
$10^{-9.0}$	−1.1	−7.0	−2.8	−1.2	3.0	− .1	3.8	3.8
$10^{-9.5}$	1.5	− .6	.6	1.1	− .8	−1.3	.4	.4

It can be seen in Table III that the observations for both species are badly fitted by the "most probable number" theory (columns 2) and that the departures from expectation do not appear to be random. The fits are sufficiently bad to suggest that the assumptions underlying the theory are violated. It was shown that the "most probable number" may be interpreted as $y(1 - q)$, where y is the number of organisms at a specified dilution and q is a constant in the expression $1 - q^x$ which

relates the response of the animal to the number of organisms it received. Unfortunately this makes it impossible to make any statement about the number of organisms necessary to produce response in an animal. The bad fits apparently make it necessary to abandon the expression $1 - q^x$, at least in this example.

Table II* shows that the fifty per cent points determined by the other methods are much alike, both between and within species. The departures from expectation (columns 1, 3, 4 of Table III) are tolerable, at least for Species A, and are sufficiently alike to show that the fitted values, P_{d_i} , from the three theories are close together. Because the results are alike and the assumptions under which they were obtained are different, it is again impossible to make any statement about the number of organisms involved. It does seem fair to state that this experiment has shown no differences in the responses of the species to the organisms.

This example suggests that observations of this nature may be fitted by curves resulting from discordant sets of assumptions but that the comparison between species may be valid. However, if one is interested in the number of organisms per *se*, it would seem to be desirable to use a more direct approach than statistical theory to determine it.

REFERENCES

1. Wilson, E. B. and Worcester, J. (1943). Bio-assay on a General Curve. *Proc. Nat. Ac. Sci.*, *29*, 150–154.
2. Berkson, J. (1951). Relative Precision of Minimum Chi-Square and Maximum Likelihood Estimates of Regression Coefficients. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, 471–479. University of California Press.
3. Cochran, W. G. (1950). Estimation of Bacterial Densities by Means of the "Most Probable Number". Biometrics *6*, 105–116.
4. Fuller, H. S. (1953). Studies of Human Body Lice, Pediculus Humanus Corporis. II. Quantitative Comparison of the Susceptibility of Human Body Lice and Cotton Rats to Experimental Infection with Epidemic Typhus Rickettsiae. *Am. J. Hyg. 58*, 188–206.

*It is of some interest to note that in this example the constants c and y in the expression

$$Q_{d_i} = c\left[1 - \frac{1}{d_i y}(1 - e^{-d_i y})\right]$$

were found to be more highly correlated than the constants α and γ in the logistic. This correlation, combined with an unfortunate choice of trial values, made the solution of the maximum likelihood equations extremely tedious.

THE ANALYSIS OF VARIANCE OF DIALLEL TABLES

B. I. Hayman

A.R.C. Unit of Biometrical Genetics, Department of Genetics, University of Birmingham

1. *Introduction*

A diallel cross is the set of n^2 possible single crosses and selfs between n homozygous (inbred) lines; it provides a powerful method of investigating the relative genetical properties of these lines. A diallel table is a set of n^2 measurements associated with a diallel cross, e.g. measurements from the progeny of a diallel cross, or from later generations obtained by selfing or backcrossing these progeny. A summary of a method of describing the genetical situation generating a diallel table has already appeared (Jinks and Hayman, 1953) and fuller accounts will appear in papers by Jinks and by Hayman. Here an analysis of variance is described which tests additive and dominance effects in diallel tables obtained from the progeny of a diallel cross.

2. *Additive systems*

A single diallel table will be considered at first, but in practice it is desirable to replicate the experiment to provide estimates of error from the block interactions, because many of even the more complex interactions within the diallel table have a genetical meaning. Suppose that the measured character is controlled by genes at k loci. In the simplest genetical system with the genes acting independently and additively the measurement of the progeny of a single cross is the mean of the two parental measurements. Maternal effects may cause differences between the progeny of reciprocal crosses so that we suppose the additive property to hold for means of reciprocal crosses. Let y_{rs} be the entry in the rth row and sth column of the diallel table, the common parent of each row being of one sex, and the common parent of each column of the other sex. (Hermaphrodites would be used as male

parents for rows and as female parents for columns). The appropriate statistical model to test for additive variation between the parents, and for maternal effects, is obtained by fitting constants to the table as follows

$$y_{rs} = m + j_r + j_s + j_{rs} + k_r - k_s + k_{rs}$$

where m = grand mean,
j_r = mean deviation from the grand mean due to the rth parents,
j_{rs} = remaining discrepancy in the rsth reciprocal sum,
$2k_r$ = difference between the effects of the rth parental line used as male parent and as female parent,
$2k_{rs}$ = remaining discrepancy in the rsth reciprocal difference.

Table 1 is the corresponding analysis of variance. A dot indicates summation over all values from 1 to n of the omitted suffix and the

TABLE 1.

	Constant	Sum of Squares	Degrees of freedom
a	j_r	$\Sigma (y_{r.} + y_{.r})^2/2n - 2y_{..}^2/n^2$	$n - 1$
b	j_{rs}	$\Sigma (y_{rs} + y_{sr})^2/4 - \Sigma (y_{r.} + y_{.r})^2/2n + y_{..}^2/n^2$	$\frac{1}{2}n(n-1)$
c	k_r	$\Sigma (y_{r.} - y_{.r})^2/2n$	$n - 1$
d	k_{rs}	$\Sigma (y_{rs} - y_{sr})^2/4 - \Sigma (y_{r.} - y_{.r})^2/2n$	$\frac{1}{2}(n-1)(n-2)$
Total		$\Sigma y_{rs}^2 - y_{..}^2/n^2$	$n^2 - 1$

sigmas summation over all values of r or r and s. The four sums of squares measure

(a) variation between the mean effects of each parental line,
(b) variation in the reciprocal sums not ascribable to (a),
(c) average maternal effects of each parental line,
(d) variation in the reciprocal differences not ascribable to (c).

This analysis was given by Yates (1947) who used (b) as the error against which to test line differences (a), and (d) as the error for maternal effects (c). That is equivalent to analysing separately the row (or column) means of two distinct two-way tables, one containing the sums of measurements from reciprocal single crosses, and the other the differences of reciprocals.

3. *Dominance*

The inclusion of dominance in the genetical system alters the situation radically. Since the deviation of progeny from their parental mean depends on dominance, (b) in Table 1 is a measure of dominance. Hence, in the absence of replication, (d) must be used as the common error against which to test (a), (b) and (c).

To interpret the components (a) and (b) more precisely we introduce a biometrical genetical model similar to Mather's (1949) specification of the effects of a polygenic system. As there are n (> 2) homozygous parents in a diallel cross we consider multiple allelic systems and suppose that m_i different alleles occur at the ith locus ($i = 1, 2, \cdots k$) in the set of parents. The genotype at the ith locus of any individual may be represented by a pair of integers (a, b) where a and $b = 1, 2, \cdots m_i$. The whole genotype controlling a character is represented by k pairs of numbers (a, b). In a parent the representation is k pairs of identical numbers (a, a).

If the genes at non-homologous loci do not interact let d_{abi} be the contribution of (a, b) at the ith locus to the measurement. Then the measurements of two parents and their F_1 are respectively $\sum_i d_{ai}$, $\sum_i d_{bi}$ and $\sum_i d_{abi}$ (writing d_{ai} for d_{aai}). In the additive system of section 2, $d_{abi} = \frac{1}{2}(d_{ai} + d_{bi})$ but, with interaction between alleles at homologous loci, i.e. dominance, we put $d_{abi} = h_{abi} + \frac{1}{2}(d_{ai} + d_{bi})$, h_{abi} being the measure of dominance. Lastly, let u_{ai} ($\sum_a u_{ai} = 1$) be the frequency of allele a at the ith locus in the parents.

Assuming that the genes at different loci are distributed independently in the parents, we find that the mean squares corresponding to (a) and (b) are $2n \sum_i \sum_a u_{ai} (\frac{1}{2}d_{ai} - \frac{1}{2} \sum_b u_{bi}d_{bi} + \sum_b u_{bi}h_{abi} - \sum_{b,c} u_{bi}u_{ci}h_{bci})^2 + \sigma_e^2$ and $2 \sum_i \sum_{a,b} u_{ai}u_{bi} (h_{abi} - \sum_c u_{ci}h_{aci} - \sum_c u_{ci}h_{bci} + \sum_{c,d} u_{ci}u_{di}h_{cdi})^2 + \sigma_e^2$. σ_e^2 is the variance of entries in the diallel table due to environmental causes and is assumed to be independent of the genetic variation. Table 3 contains in the second column the corresponding quantities for the two-allele case with $u_{1i} = u_i$, $u_{2i} = v_i$, $u_i - v_i = w_i$, $d_{1i} = d_i$, $d_{2i} = -d_i$ and $h_{12i} = h_i$. This is Mather's (1949, p. 74) notation, and equivalents in terms of his random mating D and H are in the fourth column. The third column contains equivalents in the notation of Jinks and Hayman (1953) with the additional definition $h = 4 \sum_i u_i v_i h_i$. We will continue to discuss the general case but essentially the same conclusions may be drawn from the simpler two-allele system.

Since (*b*) reduces to σ_e^2 only when all $h_{abi} = 0$, it clearly detects mean square dominance. The other mean square (a), which in section 2

detected additive variation, here detects dominance variation as well, unless the frequencies u_{ai} satisfy the symmetry condition given later. (a) and (b) respectively measure general and specific combining ability differences as defined by Henderson (1952). The mean squares (c) and (d) both estimate σ_e^2 in the absence of maternal effects.

At this stage biometrical genetics tends to diverge from this simple statistical approach. The obvious estimator of purely additive genetic variation is the variance of the parental measurements—the diagonal entries in the diallel table. This is $\sum_i \sum_a u_{ai} (d_{ai} - \sum_b u_{bi} d_{bi})^2$ (Jinks and Hayman D in the two-allele case), whether or not dominance is present, but unfortunately we cannot test its significance by this analysis of variance. Many other interesting statistics exist whose significance is difficult to establish.

However, we can extend the linear statistical model of section 2 by fitting constants for the dominance difference between parental mean and progeny mean and for deviations from this due to specific parents. The new corresponding sums of squares will be components of (b) but their meaning may not be clear until they have been expressed in terms of genetical parameters. Let

$$y_{rs} = m + j_r + j_s + l + l_r + l_s + l_{rs} + k_r - k_s + k_{rs} \qquad (r \neq s)$$

$$y_r = m + 2j_r - (n-1)l - (n-2)l_r \qquad (\text{for } y_{rr})$$

The new constants are

l = mean dominance deviation,
l_r = further dominance deviation due to the rth parent,
l_{rs} = remaining discrepancy in the rsth reciprocal sum.

The sum of squares (b) in Table 1 is replaced by those in Table 2. The third item is more conveniently obtained as a difference.

TABLE 2.

	Constant	Sum of Squares	Degrees of freedom
b_1	l	$(y_{..} - ny_{.})^2/n^2(n-1)$	1
b_2	l_r	$\Sigma (y_{r.} + y_{.r} - ny_r)^2/n(n-2) - (2y_{..} - ny_{.})^2/n^2(n-2)$	$n-1$
b_3	l_{rs}	$\Sigma (y_{rs} + y_{sr})^2/4 - \Sigma y_r^2 - \Sigma (y_{r.} + y_{.r} - 2y_r)^2/2(n-2) + (y_{..} - y_{.})^2/(n-1)(n-2)$	$\frac{1}{2}n(n-3)$

In terms of the biometrical genetical model the mean square (b_1) is $n^2 (\sum_i \sum_{a,b} u_{ai} u_{bi} h_{abi})^2/(n-1) + \sigma_e^2$ which estimates the square of

the mean dominance as expected. Table 3 contains the corresponding mean square for the two-allele case. Since h_{abi} may be either positive or negative this mean dominance may be zero without the mean square dominance vanishing. The mean square (b_2) is $4n \sum_i \sum_a u_{ai} (\sum_b u_{bi} h_{abi} - \sum_{b,c} u_{bi} u_{ci} h_{bci})^2/(n-2) + \sigma_e^2$. This reduces to σ_e^2 either in the absence of dominance or when the gene frequencies satisfy the symmetry

TABLE 3.

	Mean squares with two alleles	Jinks and Hayman (1953)	Mather (1949)
a	$2n \Sigma u_i v_i (d_i - h_i w_i)^2 + \sigma_e^2$	$\frac{1}{2}n(D - F + H_1 - H_2) + E$	$\frac{1}{2}nD + E$
b	$8 \Sigma u_i^2 v_i^2 h_i^2 + \sigma_e^2$	$\frac{1}{2}H_2 + E$	$\frac{1}{2}H + E$
b_1	$4n^2(\Sigma u_i v_i h_i)^2/(n-1) + \sigma_e^2$	$\frac{1}{4}n^2h^2/(n-1) + E$	
b_2	$4n \Sigma u_i v_i w_i^2 h_i^2/(n-2) + \sigma_e^2$	$n(H_1 - H_2)/(n-2) + E$	
c	σ_e^2	E	E
d	σ_e^2	E	E

relation $u_{ai} = \sum_b H_{abi} / \sum_{a,b} H_{abi}$ where H_{abi} is the cofactor of h_{abi} in the determinant $\{h_{abi}\}$ $(a, b = 1, 2, \cdots m_i)$. This is also the condition that mean square (a) should detect only additive variation. Illustrative examples of this relation are

(i) When h_{abi} = constant for all $a \neq b$ then $u_{ai} = 1/m_i$, i.e. all alleles at any one locus are equally frequent.

(ii) In the two-allele case $u_i = v_i = \frac{1}{2}$, which is also obvious from Table 3.

(iii) In the three-allele case u_{1i} ; $u_{2i} : u_{3i} = h_{23i} (h_{31i} + h_{12i} - h_{23i}) : h_{31i} (h_{12i} + h_{23i} - h_{31i}) : h_{12i} (h_{23i} + h_{31i} - h_{12i})$.

The mean square (b_3) also estimates dominance but has no simple interpretation, though, when the gene frequencies satisfy the symmetry relation, (b_1) and (b_3) together provide a test of dominance equivalent to (b).

4. *Subdividing the experiment*

Limitations of labour and equipment may necessitate the performance of the diallel cross in sections in different places or at different times, as in a *Drosophila* experiment of Durrant and Mather (1954). If a Latin square is superimposed upon the diallel table each letter indicates

a set of single crosses which may be performed apart from the other sets. In the analysis of variance the sum of squares for the time or distance effect is computed in the way usual for Latin squares. The letters of the Latin square are orthogonal to its rows and columns so that this sum of squares is independent of (a) and (c) in the analysis of variance of the diallel table; it is not independent of the other components. The analysis of variance thus contains the time component, (a), (c) and a remainder.

By restricting the Latin square, further orthogonal items may be extracted from the above remainder sum of squares. If the Latin square is symmetrical about the main diagonal of the diallel table, i.e. each pair of reciprocal crosses lies in one set, then (d) is also independent of the time effect. When the Latin square has n different letters in the leading diagonal, so that each self lies in a different set, (b_1), the measure of mean dominance, is orthogonal to the time effect. When n is odd the Latin square can both be symmetrical and have n different letters in the diagonal and an analysis is possible into the independent sums of squares (a), (b_1), (c), (d), time effect and remainder. Unfortunately no test of mean square dominance seems possible. The second square in Table 4 is an example of a restricted 5×5 Latin square derived from

TABLE 4.

A	*B*	*C*	*D*	*E*
B	*C*	*D*	*E*	*A*
C	*D*	*E*	*A*	*B*
D	*E*	*A*	*B*	*C*
E	*A*	*B*	*C*	*D*

A	*B*	*C*	*D*	*E*
B	*D*	*E*	*C*	*A*
C	*E*	*B*	*A*	*D*
D	*C*	*A*	*E*	*B*
E	*A*	*D*	*B*	*C*

the first square by simultaneous permutation of the rows and columns at random.

5. *Worked example.*

The data used to illustrate the analysis of sections 2 and 3 were kindly supplied by Dr. Jinks. They are the flowering times, in days from a certain date in 1951, of *Nicotiana rustica* plants from a diallel cross of eight inbred varieties. These plants were grown in two blocks, each containing 64 plots; each cross or self was represented by 10 progeny, grown in two plots of 5, with one plot in each block. This duplication of the experiment provides independent tests of the significance of every one of the components described in the analysis of variance of a single diallel table. The two diallel tables, I and II, in Table 5 contain 10 times the mean flowering time per plot.

TABLE 5.

♀

I	1	2	3	4	5	6	7	8	$y_{r.}$	
1	276	156	322	250	162	193	222	152	1733	
2	136	166	164	134	102	150	96	90	1038	
3	246	158	416	213	160	222	128	166	1709	
♂ 4	318	132	218	272	138	195	108	124	1505	
5	150	124	164	164	156	158	100	114	1130	
6	182	136	204	216	133	174	112	120	1277	
7	174	86	194	142	86	92	58	94	926	
8	152	128	158	136	126	114	84	142	1040	
$y_{.r}$	1634	1086	1840	1527	1063	1298	908	1002	10358	$y_{..}$
$y_{r.} + y_{.r}$	3367	2124	3549	3032	2193	2575	1834	2042	20716	$2y_{..}$
$y_{r.} - y_{.r}$	99	−48	−131	−22	67	−21	18	38	1660	$y_{.}$
$y_{r.} + y_{.r} - 8y_{r}$	1159	796	221	856	945	1183	1370	906	7436	$2y_{..} - 8y_{.}$
$y_{rs} - y_{sr}$		−20	−76	68	−12	−11	−48	0		
			−6	−2	22	−14	−10	38		
				5	4	−18	66	−8		
					26	21	34	12		
						−25	−14	12		
							−20	−6		
								−10		

♀

II	1	2	3	4	5	6	7	8	$y_{r.}$	
1	302	178	274	246	140	204	254	154	1752	
2	142	175	136	128	128	174	116	114	1113	
3	242	174	360	178	140	208	160	154	1616	
♂ 4	204	138	206	210	130	192	138	176	1394	
5	180	140	156	146	176	192	104	170	1264	
6	186	146	202	222	150	166	136	176	1384	
7	162	100	162	100	98	84	48	142	896	
8	154	138	140	144	124	112	96	166	1074	
$y_{.r}$	1572	1189	1636	1374	1086	1332	1052	1252	10493	$y_{..}$
$y_{r.} + y_{.r}$	3324	2302	3252	2768	2350	2716	1948	2326	20986	$2y_{..}$
$y_{r.} - y_{.r}$	180	−76	−20	20	178	52	−156	−178	1603	$y_{.}$
$y_{r.} + y_{.r} - 8y_{r}$	908	902	372	1088	942	1388	1564	998	8162	$2y_{..} - 8y_{.}$
$y_{rs} - y_{sr}$		−36	−32	−42	40	−18	−92	0		
			38	10	12	−28	−16	24		
				28	16	−6	2	−14		
					16	30	−38	−32		
						−42	−6	−46		
							−52	−64		
								−46		

The computations should be carefully arranged as in Table 5. Diallel table III contains the sum of corresponding pairs of entries in the first two diallel tables. Beside each of the three diallel tables are the row sums $y_{r.}$ and below them are the column sums $y_{.r}$, the combined

TABLE B Cont.

				♀					
III	1	2	3	4	5	6	7	8	$y_{r.}$
1	578	334	596	496	302	397	476	306	3485
2	278	341	300	262	230	324	212	204	2151
3	488	332	776	391	300	430	288	320	3325
♂ 4	522	270	424	482	268	387	246	300	2899
5	330	264	320	310	332	350	204	284	2394
6	368	282	406	438	283	340	248	296	2661
7	336	186	356	242	184	176	106	236	1822
8	306	266	298	280	250	226	180	308	2114
$y_{.r}$	3206	2275	3476	2901	2149	2630	1960	2254	20851 $y_{..}$
$y_{r.} + y_{.r}$	6691	4426	6801	5800	4543	5291	3782	4368	41702 $2y_{..}$
$y_{r.} - y_{.r}$	279	−124	−151	−2	245	31	−138	−140	3263 $y_{.}$
$y_{r.} + y_{.r} - 8y$	2067	1698	593	1944	1887	2571	2934	1904	15598 $2y_{..} - 8y_{.}$
$y_{rs} - y_{sr}$		−56	−108	26	28	−29	−140	0	
			32	8	34	−42	−26	62	
				33	20	−24	68	−22	
					42	51	−4	−20	
						−67	−20	−34	
							−72	−70	
								−56	

row and column sums $y_{r.} + y_{.r}$, the row and column differences $y_{r.} - y_{.r}$, the parental deviations $y_{r.} + y_{.r} - ny_{r}$ and the full set of differences between reciprocal crosses. The totals of the sets of sub-totals provide simple checks and the values of $y_{..}$ and $2y_{..} - ny_{.}$. The parental totals $y_{.}$ have been placed at the ends of the rows of values of $y_{r.} - y_{.r}$ which, of course, sum to zero.

Table 6 contains intermediate sums of squares, computed directly for the first two diallel tables, and halved for the third. The formulae

TABLE 6.

	I	II	III
Σy_{rs}^2	1,931,932	1,890,133	3,795,662
$y_{..}^2/n^2$	1,676,378	1,720,360	3,396,595
$\Sigma\ (y_{r.} + y_{.r})^2/2n$	3,538,105	3,543,103	7,070,907
$(y_{..} - ny_{.})^2/n^2(n-1)$	19,058	12,128	30,797
$\Sigma\ (y_{r.} + y_{.r} - ny_{r})^2/n(n-2)$	161,432	192,004	350,946
$(2y_{..} - ny_{.})^2/n^2(n-2)$	143,995	173,485	316,794
$\Sigma\ (y_{r.} - y_{.r})^2/2n$	2,278	8,086	6,739
$\Sigma\ (y_{rs} - y_{sr})^2/4$	12,168	17,754	19,112

of Tables 1 and 2, applied to the third intermediate set of sums of squares provide the final sums of squares (a), (b_1), (b_2), (b_3), (c) and (d) which measure mean effects over the two diallel tables. The excesses of the totals of the two similar final sums of squares for the first two diallel tables over the final sums for the third measure the block interactions or errors of the mean effects. As a check, (b_3) and its block interaction can be computed from sums of reciprocal crosses, but we have simply obtained them by difference from the total sums of squares. The sum of squares (B) for the overall block difference is computed in the usual way. Table 7 contains this analysis of variance. (b) is the sum of

TABLE 7.

	Sum of Squares	df	Mean square	*P*
a	277;717	7	39,674	<.001
b_1	30,797	1	30,797	<.001
b_2	34,153	7	4,879	<.001
b_3	37,289	20	1,864	<.001
b	102,238	28	3,651	<.001
c	6,739	7	963	.05–.01
d	12,373	21	589	.20–.10
t	399,067	63		
B	142	1	142	—
Ba	10,016	7	1,431	
Bb_1	390	1	390	
Bb_2	1,803	7	258	
Bb_3	3,241	20	162	
Bc	3,625	7	518	
Bd	7,185	21	342	
Bt	26,260	63	417	
Total	425,470	127		

(b_1), (b_2) and (b_3), (t) is the sum of the main effects apart from (B), and (Bt) is the sum of the interaction sums of squares.

Each error is the interaction with the environment of the corresponding mean effect and, since we would not expect, for example, additive and dominance variation to be influenced to the same extent by the environment, we must generally test each mean effect against its own

interaction. However, Bartlett's test for heterogeneity of the six error variances gives $\chi^2_5 = 6.4$, so that in this case the error variances may be pooled to give (Bt) as a common error variance. Comparison with this provides the significance levels in the last column of Table 7.

The interpretation of the results is straightforward. The significance of (a) shows genetical variation amongst the parents and of (b) dominance at some of the loci. The parental mean is greater than the progeny mean (from (b_1)) indicating dominance for early flowering time. The significance of (b_2) implies asymmetry in the gene distribution. The two items (c) and (d) show that some maternal effect may be present. Finally, there is no evidence that the difference in environment between the blocks (B) has caused any variation in flowering time.

6. *Summary*

An anlaysis of variance of diallel tables is developed which detects both additive genetic variation and dominance deviations. The mean squares are formulated in terms of a biometrical genetical model. Flowering times from a diallel cross of eight inbred varieties of *Nicotiana rustica* are analysed and the type of genetic variation present described.

7. *References*

Durrant, A. and Mather, K. Heritable variation in a long inbred line of *Drosophila*, *Genetica*, in the press.

Hayman, B. I. The theory and analysis of diallel crosses, *Genetics*, in the press.

Henderson, C. R. Specific and general combining ability, *Heterosis*, Iowa State College Press, 352–370, 1952.

Jinks, J. L. and Hayman, B. I. The analysis of diallel crosses, *Maize Genetics Cooperation News Letter*, *27*, 48–54, 1953.

Jinks, J. L. The analysis of heritable variation in a diallel cross of *Nicotiana rustica* varieties. *Genetics*, in the press.

Mather, K. *Biometrical Genetics*, London, Methuen and Co., 1949.

Yates, F. Analysis of data from all possible reciprocal crosses between a set of parental lines. *Heredity*, *1*, 287–301, 1947.

A CONFIDENCE INTERVAL FOR A PERCENTAGE INCREASE

Irwin Bross

Cornell University Medical College

INTRODUCTION

In the various statistical fields, and particularly in economic and biological situations, it is often necessary to compare the proportion of individuals in a specified class in samples from two populations. For example we may want to compare the accident rate in a plant for two consecutive months, or we may wish to contrast the proportion of individuals with mental health problems in two economic or ethnic groups, or we may want to present evidence of the reduction in the scrap-rate of a manufactured item. One index which is used in the presentation of this type of information is the percentage increase (or decrease). Thus we might speak of a 50% increase in the accident rate from one time period to another, or we might note that one group had a 20% higher incidence of mental health problems.

The percentage increase is often a useful index because it boils the information down to a single number and this number is rather readily interpreted. On the other hand the percentage increase may be very misleading when small proportions are involved and this may be true even when the sample sizes are large. Thus if there are six accidents in one month and nine accidents in the next month we might state that "There is a 50% increase in the accident rate." This rather frightening statement is quite misleading because on the basis of the two figures there is no strong evidence that any *real* change in the accident rate has occurred.

One way to avoid misinterpretation of percentage increases is to present a *confidence interval* for the percentage increase instead of the single number. If this confidence interval is very wide, then this will serve as a warning that the estimate of the percentage increase should not be interpreted too literally. In the next section instructions will be given for computing an exact confidence interval for a percentage increase which applies in the important special case where small pro-

portions are involved. In the last section a justification for the procedures will be presented.

CALCULATION OF CONFIDENCE INTERVALS

In the summaries of articles in the medical literature one often encounters statements such as: "It was found that there were 40% more complications when Drug B was used than when Drug A was used."

The body of such papers may provide the actual data on the comparison of a standard drug, A, with a new drug, B, and these data may resemble the following artificial data:

Treatment	Total Number of Cases	Number of Cases with Complications	Percentage of Complications
Drug A	189	13	6.88
Drug B	104	10	9.62
TOTALS	293	23	7.85

The percentage increase in complications with Drug B is:

$$\hat{\theta} = 100\,\frac{9.62 - 6.88}{6.88} = 100\,\frac{2.74}{6.88} = 39.8\%$$

so that the summary statement is numerically correct. It is, however, a rather misleading statement and the calculation of a confidence interval brings this point out clearly.

To calculate a 95% confidence interval for the percentage increase we first calculate a 95% confidence interval for P, the ratio of cases with complications on Drug A to total cases with complications. This latter confidence interval is readily obtained, for it is merely the usual binomial confidence interval. The confidence interval thus obtained serves only as a useful computational device.

The simplest way to determine binomial confidence intervals is to use tables or charts such as the ones found in the appendix of Dr. Mainland's *Elementary Medical Statistics* (1).

In Mainland's notation two quantities are necessary to use his table, "N" and "The number of A's." For our problem N is the total number of cases with complications (i.e., the total number of individuals in the specified class), so $N = 23$. The quantity N is the sum of the respective numbers in the two samples and the *smaller* of the two numbers is "the number of A's." If this number is associated with the population used

as the *base* of the percentage increase, then the tables give the confidence limits for P. Otherwise the tables give the limits for $Q = 1 - P$ and the limits for P are found by subtraction.

Thus in the above example "the number of A's" equals 10, the number of complications in the patients given Drug B. Hence the tabular confidence limits, .232 and .655, must be subtracted from one to give:

Lower Limit (L.L.) = .345 Upper Limit (U.L.) = .768

These limits are used to calculate the confidence interval for the percentage increase. The formulas used below will be derived in the next section. The lower limit for the percentage increase, L, is found from:

$$\mathrm{L} = 100\,\frac{n_1 - (n_1 + n_2)(\mathrm{U.L.})}{n_2(\mathrm{U.L.})} = 100\,\frac{189 - 293(.768)}{104(.768)}$$

$$L = 100\,\frac{-36.02}{79.87} = -45\%$$

where n_1 and n_2 are the two sample sizes.

The upper limit is found from:

$$U = 100\,\frac{n_1 - (n_1 + n_2)(\mathrm{L.L.})}{n_2(\mathrm{L.L.})} = 100\,\frac{189 - 293(.345)}{104(.345)}$$

$$\mathrm{U} = 100\,\frac{87.92}{35.88} = 245\%$$

Hence the 95% confidence interval for the percentage increase is from $-$ 45% to 245%. When these limits are presented it is evident that we do not have a very good idea of the magnitude of the percentage increase. In fact we are not even confident that there *is* an increase.

The confidence limits therefore serve to mitigate the misleading impression that is conveyed by the statement: "There is a 40% increase in complications with Drug B."

NOTE: If a table of binomial confidence limits is not available, normal approximations may be used:

$$\text{Approx. L.L.} = \hat{P} - 2\sqrt{\frac{\hat{P}\hat{Q}}{N}}$$

$$\text{Approx. U.L.} = \hat{P} + 2\sqrt{\frac{\hat{P}\hat{Q}}{N}}$$

where $\hat{P}$ is the ratio of the number of individuals in the specified class in the reference group to the total number of individuals in the specified class. Here $\hat{P} = 13/23 = .565$, and:

$$\text{Approx. L.L.} = .565 - 2\sqrt{\frac{(.565)(.435)}{23}} = .358$$

$$\text{Approx. U.L.} = .565 + 2\sqrt{\frac{(.565)(.435)}{23}} = .772$$

which give as approximate limits for the percentage increase:

$$\text{Approx. L} = -46\% \qquad \text{and} \qquad \text{Approx. U} = 226\%$$

JUSTIFICATION

The confidence limits for percentage increase presented here are derived under the assumption that the proportion of individuals in the specified class is small in both samples and consequently the number of individuals in the specified class will follow the Poisson distribution.

Let the true proportion of individuals in the specified class be p_1 and p_2 in the two respective populations. Thus the real percentage increase, θ, will be defined by (1.01).

$$\theta = 100\,\frac{p_2 - p_1}{p_1} \tag{1.01}$$

Let x_1 be the number of individuals in the designated class in a sample of size n_1 from the first population, and let x_2 and n_2 be the corresponding quantities for the second population. We wish to use x_1 , x_2 , n_1 , and n_2 to obtain a confidence interval for the percentage increase, θ. The ordinary estimate of the percentage increase would be:

$$\hat{\theta} = 100\,\frac{\frac{x_2}{n_2} - \frac{x_1}{n_1}}{\frac{x_1}{n_1}} = 100\,\frac{n_1x_2 - n_2x_1}{n_2x_1} \tag{1.02}$$

which in the case where the sample sizes are equal reduces to:

$$\theta = 100\,\frac{x_2 - x_1}{x_1} \tag{1.03}$$

The confidence interval is easily obtained by starting with the fact that x_1 and x_2 follow the Poisson distribution and hence:

$$P(x_1 , x_2) = \frac{e^{-n_1p_1}(n_1p_1)^{x_1}}{x_1!}\,\frac{e^{-n_2p_2}(n_2p_2)^{x_2}}{x_2!} \tag{1.04}$$

If equation (1.04) is rewritten in the form (1.05), the solution of the problem is immediate.

$$(1.05) \qquad P(x_1 , x_2) = \frac{e^{-(n_1p_1+n_2p_2)}(n_1p_1 + n_2p_2)^{x_1+x_2}}{(x_1 + x_2)!} \cdot \frac{(x_1 + x_2)!}{x_1!x_2!}\left(\frac{n_1p_1}{n_1p_1 + n_2p_2}\right)^{x_1}\left(\frac{n_2p_2}{n_1p_1 + n_2p_2}\right)^{x_2}$$

In form (1.05) the probability distribution of x_1 and x_2 had been written as the product of a Poisson distribution and a binomial distribution. These two distributions correspond to $P(x_1 + x_2)$ and $P(x_1 \mid x_1 + x_2)$ respectively, i.e.,

$$P(x_1 , x_2) = P(x_1 + x_2)P(x_1 \mid x_1 + x_2)$$

If $x_1 + x_2$ is regarded as fixed, then the distribution of x_1 and x_2 follows the ordinary binomial.

$$(1.06) \qquad P(x_1 \mid x_1 + x_2) = \frac{(x_1 + x_2)!}{x_1!x_2!} P^{x_1}Q^{x_2}$$

where

$$P = \frac{n_1p_1}{n_1p_1 + n_2p_2} \qquad \text{and} \qquad Q = \frac{n_2p_2}{n_1p_1 + n_2p_2}$$

Therefore x_1 , x_2 , n_1 , and n_2 can be used to put confidence limits on P by applying the well known procedures for binomial confidence limits. Moreover since P is monotonically related to θ:

$$(1.07) \qquad \theta = 100\,\frac{n_1 - (n_1 + n_2)P}{n_2P}$$

This leads at once to confidence limits on θ.

To see how this device can be applied, suppose that the usual methods are employed to give a 95% confidence interval on P. If L.L. and U.L. are respectively the lower and upper 95% endpoints, then by the definition of a confidence interval:

$$P(\text{L.L.} \leq P \leq \text{U.L.}) = .95$$

By a well known theorem, if $f(P)$ is a monotonically decreasing function of P, then:

$$(1.08) \qquad P\{f(\text{U.L.}) \leq f(P) \leq f(\text{L.L.})\} = .95$$

It is easy to show that θ is a monotonically decreasing function of P, since the derivative of θ with respect to P is $-100n_1/n_2P^2$. Therefore

it follows from (1.07) and (1.08) that:

$$P\{L \leq \theta \leq U\} = .95 \tag{1.09}$$

where L and U represent the lower and upper limits of the confidence interval:

$$L = 100\,\frac{n_1 - (n_1 + n_2)(U.L.)}{n_2(U.L.)} \qquad U = 100\,\frac{n_1 - (n_1 + n_2)(L.L.)}{n_2(L.L.)} \tag{1.10}$$

For the special case where $n_1 = n_2$ the limits simplify to:

$$L = 100\,\frac{1 - 2(U.L.)}{(U.L.)} \qquad U = 100\,\frac{1 - 2(L.L.)}{(L.L.)} \tag{1.11}$$

Note that in (1.10) it is not necessary to know n_1 and n_2 , but only the ratio n_1/n_2 . This is fortunate since in some practical applications n_1 and n_2 may not be known. For example in presenting accident statistics the number of persons injured (x_1 and x_2) will be known, but the number of individuals exposed to risk (n_1 and n_2) may not be known. However it may be known that the number exposed to risk is about the same or that there are twice as many in one population as in the other and this sort of information will be enough to establish a confidence interval.

REFERENCES

(1) Mainland, D., *Elementary Medical Statistics*, 1952.

CHAIN BLOCK DESIGNS WITH TWO-WAY ELIMINATION OF HETEROGENEITY

JOHN MANDEL
National Bureau of Standards, Washington, D. C.

1. INTRODUCTION

In a recent paper, Youden and Connor [3] presented a new class of experimental designs, which they call chain block designs. Their paper contains formulas for the estimation of the treatments corrected for blocks, for the blocks corrected for treatments, and more generally, for the construction of the analysis of variance appropriate for these designs. In the present paper a particular class of these designs is generalized in such a fashion that the elimination of bias is achieved not only for blocks, but also for another factor, which may, for example, be identified with order within blocks. It will be convenient to identify the blocks with the columns, and the second factor with the rows of a rectangular pattern. The designs presented in this paper have a similar relation to a class of chain block designs as have the Youden squares to the balanced incomplete blocks. For brevity's sake, the new designs will frequently be referred to, in this paper, as "generalized chain-blocks". The flexibility of the new designs reflects that inherent in the simple chain blocks. The only restrictions for the generalized designs are that the number of blocks be even and that the number of treatments be a multiple of the number of blocks. The number of replications of each treatment is two in the basic generalized chain block, but by considering groups of treatments this restriction can be removed. The calculations involved in the analysis of the new designs are simple.

2. CHAIN BLOCK DESIGNS

a. *The Simple Chain Block*

Let α_1 , α_2 , α_3 , $\cdots$, α_v denote v treatments or groups of treatments. (In the latter case, we will assume each group to be composed of the same number of treatments.)

Youden and Connor [3] have introduced the design shown in table I for testing v treatments in v blocks. In their terminology, the table represents a chain block design, in which all treatments belong to class C_2 , that is, each treatment occurs in duplicate. The symbols a_i and a_i' represent duplicate yields for treatment α_i . It is possible to construct simple formulas for the estimation of treatments (corrected for columns and rows), of block effects (column effect), and of row effect (replication differences). These formulas, together with the analysis of variance, are given in section 4.

TABLE I.—SIMPLE CHAIN BLOCK

Replication \ Block	1	2	3	···	i	···	v
I	a_1	a_2	a_3	···	a_i	···	a_v
II	a'_2	a'_3	a'_4	···	a'_{i+1}	···	a'_1

A restrictive feature of the simple chain block with duplicate observations for each treatment is the equality of the number of blocks and the number of treatments. This restriction can be removed in two different ways:

(1) By considering groups of treatments in each cell of the chain block; (2) By distributing the treatments in more than one row.

The second generalization, made possible by the introduction of chain blocks with two-way elimination of heterogeneity, is discussed in the following section. The first generalization was considered in Youden and Connor's original paper and consists in letting the symbols α_1 , α_2 , $\cdots$, α_v stand for groups of treatments. Accordingly, the symbols a_1 , a_2 , $\cdots$, a_v , a'_1 , a'_2 , $\cdots$, a'_v represent average yields for the corresponding groups. Thus, if α_i represents a group of treatments α_{i1} , α_{i2} , α_{i3} , $\cdots$, α_{in} , then a_i will represent the average yield of the first replication of this group and a'_i the average yield of the second replication. The treatments α_{i1} , α_{i2} , $\cdots$, α_{in} of group α_i are not necessarily different. They may represent within group replications of one or more treatments. In this fashion, chain block designs are obtained in which the number of replications of some or all treatments exceeds two.

The analysis of chain block designs with grouped treatments is best carried out by first calculating, by means of the formulas given in section 4, the block "biases" on the basis of the average yields of groups, and then "correcting" each individual treatment yield by subtracting from it the bias of the block in which it occurs. The average of the corrected values of all replicates of a particular treatment is the estimate for this treatment, corrected for block effects.

b. *The Chain Block with Two-Way Elimination of Heterogeneity*

Consider q sets of treatments, such that each set consists of the same, even number of treatments, say $2t$, as indicated in table II, where each letter represents a different treatment. (Thus a_i and A_i are different treatments.)

TABLE II.—SETS OF TREATMENTS

Set 1:	a_1	a_2	⋯	a_t	A_1	A_2	⋯	A_t
Set 2:	b_1	b_2	⋯	b_t	B_1	B_2	⋯	B_t
⋯⋯⋯⋯								
Set q:	q_1	q_2	⋯	q_t	Q_1	Q_2	⋯	Q_t

We will construct a design for the comparison of the $2tq$ treatments of table II, using $2t$ blocks of $2q$ treatments each. Thus, each treatment will occur in 2 replications. This design is given in table III.

TABLE III.—GENERALIZED CHAIN BLOCKS

Row \ Block	1	2	3	⋯	t	$t+1$	$t+2$	$t+3$	⋯	$2t$
1	a_1	a_2	a_3	⋯	a_t	A_1	A_2	A_3	⋯	A_t
2	b_1	b_2	b_3	⋯	b_t	B_1	B_2	B_3	⋯	B_t
3	c_1	c_2	c_3	⋯	c_t	C_1	C_2	C_3	⋯	C_t
⋯⋯	⋯⋯					⋯⋯				
⋯⋯	⋯⋯					⋯⋯				
q	q_1	q_2	q_3	⋯	q_t	Q_1	Q_2	Q_3	⋯	Q_t
$q+1$	B_1'	B_2'	B_3'	⋯	B_t'	a_2'	a_3'	a_4'	⋯	a_1'
$q+2$	C_1'	C_2'	C_3'	⋯	C_t'	b_2'	b_3'	b_4'	⋯	b_1'
$q+3$	D_1'	D_2'	D_3'	⋯	D_t'	c_2'	c_3'	c_4'	⋯	c_1'
⋯⋯	⋯⋯					⋯⋯				
⋯⋯	⋯⋯					⋯⋯				
$2q$	A_1'	A_2'	A_3'	⋯	A_t'	q_2'	q_3'	q_4'	⋯	q_1'

The design consists essentially of 4 quadrants, the upper two being merely the $2tq$ treatments written in q rows and $2t$ columns (blocks). The lower left quadrant contains duplicate observations (indicated by primes) of the treatments occurring in the upper right quadrant with a cyclical permutation of the rows. The lower right quadrant contains duplicate observations of the treatments occurring in the upper left quadrant with a cyclical permutation of the columns. The subdivision into quadrants is made merely for convenience in the exposition and analysis of the data and has no functional meaning. Thus, each block extends over $2q$ rows and each row over $2t$ blocks. Moreover, any permutation of rows or of columns is permissible.

Now, it can be seen that the design has chain block features, both according to blocks and to rows, by grouping the treatments occurring in a single row or a single column of each quadrant. Thus, if all q

treatments occurring in any given column of each quadrant are combined into a single group, the design of table IV is obtained.

TABLE IV.

Row \ Block	1	2	⋯	t	$t+1$	$t+2$	⋯	$2t$
1 to q	z_1	z_2	⋯	z_t	Z_1	Z_2	⋯	Z_t
$q+1$ to $2q$	Z_1'	Z_2'		Z_t'	z_2'	z_3'	⋯	z_1'

The symbols z and Z stand for averages of q treatments: z_i and z_i' represent groups of the same q treatments, and so do Z_i and Z_i'.

That table IV is a simple chain block becomes immediately apparent when the table is rewritten with the blocks in the order: 1, $t+1$, 2, $t+2$, $\cdots$, t, $2t$. Consequently, the block effects can be estimated by the method given in section 4.

Similarly, by grouping, in table III, the t treatments occurring in any given row of each quadrant, one obtains the design given in table V.

TABLE V.

Row \ Block	1 to t	$t+1$ to $2t$
1	u_1	U_1
2	u_2	U_2
3	u_3	U_3
·	·	·
·	·	·
q	u_q	U_q
$q+1$	U_2'	u_1'
$q+2$	U_3'	u_2'
·	·	·
·	·	·
$2q$	U_1'	u_q'

Table V, like table IV, is a simple chain block, as is apparent by rearranging the rows in the order 1, $q+1$, 2, $q+2$, $\cdots$, q, $2q$. From this table, then, it is possible to estimate the row effects in the same way in which the block effects are estimated from table IV.

Finally, one obtains treatment estimates, free from row and block biases, by correcting each value in table III for the biases of the block and the row in which it occurs, and taking the average of the two corrected observations for each treatment. In order to clarify the

computational procedure, a numerical example is worked out in detail in section 6 of this paper.

It is interesting to note that the block effects and the row effects are estimated independently of each other; it is not necessary to correct for row effects in order to obtain the column effects, and vice versa. This property of orthogonality does not extend to the treatments; the estimates for the treatment effects depend on the estimates of both the row and the column effects. It should be noted that the analysis just outlined assumes that no interactions exist between treatments, rows, and columns. This is the usual model for incomplete block designs.

As in the case of simple chain block designs, it is, of course, possible to consider groups of treatments. In this case, each symbol, in the body of table III, would stand for the average of the yields of the treatments composing the group.

Section 4 includes a presentation of the analysis of variance of the generalized chain block design. It may be noted that all computations are simple and straighforward.

Even though the method of computation of blocks, rows and treatments described above is a plausible one, it remains to be proved that it is the correct least squares solution. This proof is given in the Appendix.

3. SOME GENERALIZED CHAIN BLOCKS

Chain blocks with two-way elimination of heterogeneity are particularly useful wherever it is required to keep the number of replicates small. In Fisher and Yates' notation [1], the block size k, the number of treatments v, the number of blocks b, and the number of replicates r, are related by the formula:

$$bk = vr$$

For any given block size k, the number of blocks necessary for testing v treatments is smallest when r is a minimum. Since r is always 2 in the basic generalized chain block, this design, when applicable, is therefore as economical as possible (barring designs without replication). A few useful designs are shown below, using Fisher and Yates' notation. The symbols used in section 2.b of this paper correspond to those of Fisher and Yates as follows:

$$k = 2q$$

$$v = 2tq$$

$$b = 2t$$

$$r = 2$$

It should be noted that, in the case of generalized chain blocks, the concepts of "blocks" (columns) and "rows" are interchangeable, so that any such design with parameters $b = b_0$ and $k = k_0$ is also a generalized chain block with parameters $b = k_0$ and $k = b_0$, the number of treatments being the same, and $r = 2$. Accordingly, any of the schemes given below, with the exception of those for which the number of rows equals the number of columns, represents two different designs.

$k = 4, v = 8$*

a	b	c	d
e	f	g	h
g	h	b	a
c	d	f	e

$k = 4, v = 12$

a	b	c	d
e	f	g	h
i	j	k	l
g	h	b	a
k	l	f	e
c	d	j	i

$k = 4, v = 16$

a	b	c	d
e	f	g	h
i	j	k	l
m	n	o	p
g	h	b	a
k	l	f	e
o	p	j	i
c	d	n	m

$k = 6, v = 18$

a	b	c	d	e	f
g	h	i	j	k	l
m	n	o	p	q	r
j	k	l	b	c	a
p	q	r	h	i	g
d	e	f	n	o	m

$k = 6, v = 24$

a	b	c	d	e	f
g	h	i	j	k	l
m	n	o	p	q	r
s	t	u	v	w	x
j	k	l	b	c	a
p	q	r	h	i	g
v	w	x	n	o	m
d	e	f	t	u	s

$k = 4, v = 20$

a	b	c	d	e	f	g	h	i	j
k	l	m	n	o	p	q	r	s	t
p	q	r	s	t	b	c	d	e	a
f	g	h	i	j	l	m	n	o	k

$k = 6, v = 30$

a	b	c	d	e	f	g	h	i	j
k	l	m	n	o	p	q	r	s	t
u	v	w	x	y	z	α	β	γ	δ
p	q	r	s	t	b	c	d	e	a
z	α	β	γ	δ	l	m	n	o	k
f	g	h	i	j	v	w	x	y	u

In using these designs it is generally advisable to use random processes for the allocation of the letters to the treatments and of the rows and columns to the corresponding variables. However, since all

*The use of the letter v in this section (number of treatments in a generalized chain block) should not be confused with the different meaning it has in section 2.a (number of blocks in a single chain block).

pairs of treatments are not compared with the same precision, a partially systematic allocation may sometimes be desirable, using the comparisons of highest precision for the comparison of those treatments in which the experimentor is mostly interested. Formulas for evaluating the precision of any treatment comparison are given in section 5.

4. CALCULATIONS FOR CHAIN BLOCK DESIGNS

a. *Simple Chain Block*

In a simple chain block each of the rows contains all the treatments once. Consequently, the estimation of row effects is at once accomplished by calculating the difference between the two row averages. If (1) $\bar{d}$ = (average yield of row I) − (average yield of row II), then the bias of row I is of course = $+\bar{d}/2$ and the bias of row II = $-\bar{d}/2$.

For the estimation of treatment effects, corrected for blocks, and of block (column) effects, corrected for treatments, the first step is to add two rows to table I, as indicated in table VI below. The entries in the added rows are defined by the following relations:

(2a) $$d_j = a_j - a'_{j+1}$$

(for $j = v$, make $j + 1 = 1$)

(2b) $$p_j = a'_{j+1} - a_{j+1}$$

(2c) $$D = \sum d_j = -\sum p_j\,; \qquad \bar{d} = \frac{D}{v}; \qquad \bar{p} = -\frac{D}{v} = -\bar{d}.$$

(2d) $$G = \text{grand sum} = (a_1 + a_2 + \cdots + a_v) + (a'_1 + a'_2 + \cdots + a'_v)$$

TABLE VI.—TREATMENT AND BLOCK EFFECTS IN SIMPLE CHAIN BLOCK

Row \ Block	1	2	3	⋯	$v-2$	$v-1$	v	Sum
I	a_1	a_2	a_3	⋯	a_{v-2}	a_{v-1}	a_v	G
II	a'_2	a'_3	a'_4	⋯	a'_{v-1}	a'_v	a'_1	
$d_j = a_j - a'_{j+1}$	d_1	d_2	d_3	⋯	d_{v-2}	d_{v-1}	d_v	D
$p_j = a'_{j+1} - a_{j+1}$	p_1	p_2	p_3	⋯	p_{v-2}	p_{v-1}	p_v	$-D$

For any value of $j (j = 1, 2, \cdots, v)$, let us denote by $\hat{\alpha}_j$ the estimate, corrected for block effects, of treatment α_j. We wish to estimate $\hat{\alpha}_k$, corresponding to a particular treatment α_k.

Consider the following sequence of v numbers, denoted as sequence S, which constitutes an arithmetic progression with common difference -2.

Sequence S: $(v-1)\quad (v-3)\quad (v-5)\quad \cdots\quad -(v-5)\quad -(v-3)\quad -(v-1)$

(For example, for $v = 10$, sequence S reads:

$$9\ 7\ 5\ 3\ 1\ -1\ -3\ -5\ -7\ -9).$$

To compute $\hat{\alpha}_k$, permute the d_i cyclically, so as to make d_k the first term, and write them below the terms of the S-sequence, as follows:

S − sequence:	$v-1$	$v-3$	$v-5$	$\cdots$	$-(v-5)$	$-(v-3)$	$-(v-1)$
d − values:	d_k	d_{k+1}	d_{k+2}	$\cdots$	d_{k-3}	d_{k-2}	d_{k-1}

Now, multiply each d-value by the associated term of the S-sequence and sum all the products. Call this sum T_k . Thus:

$$(3)\qquad T_k = (v-1)(d_k - d_{k-1}) + (v-3)(d_{k+1} - d_{k-2}) + (v-5)(d_{k+2} - d_{k-3}) + \cdots$$

Then the estimate for α_k is:

$$(4)\qquad \hat{\alpha}_k = \frac{1}{2v}(G + T_k)$$

where G is given by (2d).

The blocks, corrected for treatments, are estimated as follows: Let $e_j (j = 1, 2, \cdots, v)$ represent the bias of block j, that is, e_j is the systematic error affecting each yield occurring in block j. As usual, we will assume the sum of all the e_j to be zero, that is, all systematic errors are taken with reference to the overall average. Let $\hat{e}_j$ denote the estimate of e_j .

To compute $\hat{e}_k$, for a particular value of k, proceed first exactly as for the calculation of T_k , using p_i in the place of d_i . Denote the sum of products of the p_i with the associated terms of sequence S by B_k . Thus:

$$(5)\qquad B_k = (v-1)(p_k - p_{k-1}) + (v-3)(p_{k+1} - p_{k-2}) + (v-5)(p_{k+2} - p_{k-3}) + \cdots$$

Then, the estimate for e_k is:

$$(6)\qquad \hat{e}_k = \frac{1}{2v} B_k .$$

An alternate procedure, particularly useful for the complete analysis of a simple chain block, is as follows:

First, estimate $\hat{\alpha}_1$ and $\hat{e}_1$ as described above. Then, use the following recursion formulas for the successive estimation of $\hat{\alpha}_2, \hat{\alpha}_3, \cdots, \hat{\alpha}_v$; $\hat{e}_2, \hat{e}_3, \cdots, \hat{e}_v$.

(7a) $$\hat{\alpha}_{k+1} = \hat{\alpha}_k - d_k + \bar{d}$$

(7b) $$\hat{e}_{k+1} = \hat{e}_k - p_k + \bar{p}$$

This alternate method is particularly well adapted to computations by means of a desk calculator, since the computation of all $\hat{\alpha}$ or of all $\hat{e}$ values can be performed without clearing the machine.

The analysis of variance of simple chain blocks is given in [3]. For purposes of completeness, it is reproduced in table VII, using the notation of tables I and VI.

TABLE VII.—ANALYSIS OF VARIANCE OF SIMPLE CHAIN BLOCK

Source	Degrees of Freedom	Sum of Squares	Mean Square
Total	$2v - 1$	$S_1 = \sum a_i^2 + \sum a_i'^2 - \frac{G^2}{2v}$	MS_1
Treatments ignoring blocks	$v - 1$	$S_2 = \frac{1}{2}\sum (a_i + a_i')^2 - \frac{G^2}{2v}$	MS_2
Blocks eliminating treatments	$v - 1$	$S_3 = \frac{1}{2}\sum \hat{e}_i(p_i - p_{i-1})$	MS_3
Rows	1	$S_4 = \frac{D^2}{2v}$	MS_4

The following two identities may be used to check the computations:

$$S_1 = S_2 + S_3 + S_4$$

$$S_3 + S_4 = \tfrac{1}{2}\sum p^2 \qquad (= \text{within treatment sum of squares})$$

If it is desired to compute a sum of squares for treatments, corrected for blocks, the following formula is used:

Sum of squares for treatments, eliminating blocks,

$$= S_2'' = \tfrac{1}{2}\sum \hat{\alpha}_i(d_i - d_{i-1})$$

Since there are no degrees of freedom for error, the analysis only acquires usefulness when an independent estimate for the error mean square is available. This will be the case when the letters in table I stand for groups of treatments, or when the design is expanded to a chain block with two-way elimination of heterogeneity.

b. *Generalized Chain Block*

The method of computation for the chain block with two-way elimination of heterogeneity follows readily from the discussion in section 2b and the formulas for the simple chain block given in the preceding section.

The block (column) and row effects are estimated by applying formula (6) to tables IV and V, respectively, after appropriate reordering of the blocks in table IV and of the rows in table V. For the first table, make $v = 2t$ and for the second table, make $v = 2q$. The biases thus calculated, taken with the opposite sign, are used as additive corrections to the original observation. In doing this, mistakes are avoided by writing the corrections (biases with sign changed) near the corresponding column and row headings, as shown in the numerical example in section 6. Finally, averages are taken of the two corrected observations for each treatment.

An analysis of variance is useful to test the effectiveness of the design for the removal of block and row biases.

TABLE VIII.—ANALYSIS OF VARIANCE OF GENERALIZED CHAIN BLOCK

Source	Degrees of Freedom	Sum of Squares	Mean Square
Total	$4tq - 1$	S_1'	MS_1'
Treatments ignoring blocks and rows	$2tq - 1$	S_2'	MS_2'
Blocks eliminating treatments	$2t - 1$	S_3'	MS_3'
Rows eliminating treatments	$2q - 1$	S_4'	MS_4'
Error	$2(t - 1)(q - 1)$	S_5'	MS_5'

The sums of squares are calculated as follows, using the notation of table III.

S_1' = sum of squares of deviations of all individual observations from grand mean

$$S_2' = \tfrac{1}{2}[(a_1 + a_1')^2 + (a_2 + a_2')^2 + \cdots + (Q_t + Q_t')^2]$$

$$- \text{grand mean correction term}$$

S_3' is obtained from the simple chain block of table IV, calculating S_3 as in table VII and multiplying by q, since the observations in table IV are averages of q original observations.
S_4' is obtained from the simple chain block of table V, calculating S_3 as in table VII and multiplying by t, since the observations in table V are averages of t original observations.
S_5' is obtained by difference.

The calculations may be checked as follows:

$$S_3' + S_4' + S_5' = \tfrac{1}{2}[(a_1 - a_1')^2 + (a_2 - a_2')^2 + \cdots + (Q_t - Q_t')^2]$$

In this identity, either member represents the within treatment sum of squares. Tests of significance for blocks and rows are made by calculating the F values

$$\frac{MS_3'}{MS_5'} \quad \text{and} \quad \frac{MS_4'}{MS_5'}$$

For a test of significance of the treatment effects, the sum of squares, corrected for block and row effects, must first be calculated. This can be done by calculating sums of squares of columns (blocks) and rows, ignoring the treatments in each case, and subtracting both these sums of squares from the quantity $(S_1' - S_5')$. The remainder is the sum of squares for the corrected treatments. See section 6 for a numerical illustration.

5. PRECISION OF TREATMENT COMPARISONS

Since the chain block is not a balanced design, there will, in general, be more than one error term for the comparison of pairs of treatments. Youden and Connor [3] give the appropriate formulas for the simple chain block. In order to express the variance of the difference of two corrected treatment estimates in a generalized chain block, it is useful to introduce the following concept of "distance".

Definition: In a simple chain block, using the notation of table I, the "distance" between treatments α_i and α_k $(i \leq k)$, is the number $k - i$ or $v - (k - i)$, whichever is the smaller.

Now, consider the generalized chain block shown in table III and let V_1 and V_2 be any pair of treatments. In the construction of table IV, V_1 will occur in a z-average, say z_1 , and V_2 in z_2 . Let l represent the "distance", as defined above, between z_1 and z_2 (after reordering of the columns in table IV to form a simple chain block).

Similarly, let l' be the "distance" between the averages u_1 and u_2 , in which V_1 and V_2 respectively occur in the construction of table V.

Then the variance of the difference between the corrected estimates of V_1 and V_2 is:

$$\text{Variance } (\hat{V}_1 - \hat{V}_2) = \sigma^2\left[1 + \frac{M + M'}{qt}\right] \tag{8}$$

where σ^2 is the variance of a single observation and

$$M = \begin{cases} 0 & \text{for} \quad l = 0 \\ 2lt - t - l^2 & \text{for} \quad l \neq 0 \end{cases}$$

$$M' = \begin{cases} 0 & \text{for} \quad l' = 0 \\ 2l'q - q - l'^2 & \text{for} \quad l' \neq 0 \end{cases}$$

A sketch of the derivation of equation (8) is given in the appendix.

As an illustration, let us calculate the variance of the difference of treatments n and k in the design given in section 3 for $k = 6$, $v = 24$. We have: $t = 3$, $q = 4$. The "distance" according to columns, l, is obtained by remembering that the columns headed a, b, c, d, e, f will, after reordering, have the indices 1, 3, 5, 2, 4, 6. Since n occurs in 3, and k in 4, we have: l = either $4 - 3 = 1$, or $6 - (4 - 3) = 5$. Since $1 < 5$, $l = 1$. Similarly, one finds: $l' = 5 - 2 = 3$. Consequently:

$$\text{Variance } (\hat{n} - \hat{k}) = \sigma^2\left[1 + \frac{2 + 11}{12}\right] = \frac{25}{12}\sigma^2$$

6. A NUMERICAL EXAMPLE

In the road testing of tires for rate of treadwear, it is frequently necessary to test more tires than can be run simultaneously on one vehicle. Furthermore, the vehicle itself is not a homogeneous "block", since the treadwear of tires in different wheel positions of the same vehicle may vary several fold. Usually, as many tires are included in a test as there are wheels in all vehicles combined, and the tires are rotated among vehicles and positions from run to run, in such a way that all tires are tested equally in all positions. A number of tests have been carried out [2] in which the basic design is a latin square (generally 4×4) involving vehicles, wheel positions, test runs, and tire brands (or tire constructions) as variables. The entire test generally consists of a number of such latin squares inter-related according to a systematic pattern. The results of these tests have shown that one could obtain tire comparisons of satisfactory precision in a relatively small number of test runs, provided that it were possible to balance out the effects of wheel positions in this number of runs. This has led to the use of Youden squares and simple chain blocks in tire test designs. A further problem is encountered when it becomes desirable (as it

often does) to include in a single test more tires than can be simultaneously accommodated on the test vehicles, for example: to test 32 tires using 4 four-wheeled vehicles. In such cases it is necessary to compensate for run to run variability as well as for wheel position effects. Such double elimination of bias has been accomplished by using the generalized chain block given in section 3, for $k = 4$, $v = 8$, in lieu of the 4×4 latin square as the basic design around which the test is constructed. The columns (blocks) can be identified with the four wheel positions of a vehicle, and the rows with four test runs. The tires are the treatments. Table IX presents data obtained in a road test run on commercial tires in accordance with this design. The capital letters in the body of the table represent eight of the tires. The entire test involved 32 tires, tested in 16 runs, using 4 vehicles. The numerical values are decimal logarithms of the rates of wear. The latter are expressed in grams of rubber loss per 1000 miles. The reasons for converting the original observations to logarithms before analyzing the data are two fold. In the first place, it has been shown [2] that the experimental error of the weight loss of a tire tread tends to be proportional to the magnitude of the loss. And in the second place, differences between different tires, as well as biases due to wheel positions or to run to run effects are more truly represented by ratios than by absolute differences.

The marginal values are averages as indicated and are required for the computation of the "Position" and "Run" biases.

TABLE IX.—LOGARITHM OF RATE OF TREADWEAR

Wheel Position / Run	I Left Rear	II Right Rear	III Left Front	IV Right Front	$\frac{I + II}{2}$	$\frac{III + IV}{2}$
W	A	B	C	D	A, B	C, D
	1.802	1.862	1.173	1.762	1.8320	1.4675
X	E	F	G	H	E, F	G, H
	1.935	2.072	1.703	1.935	2.0035	1.8190
Y	G'	H'	B'	A'	G', H'	A', B'
	1.610	1.568	1.267	1.522	1.5890	1.3945
Z	C'	D'	F'	E'	C', D'	E', F'
	1.816	1.935	1.418	1.594	1.8755	1.5060
$\frac{W + X}{2}$	A, E	B, F	C, G	D, H		
	1.8685	1.9670	1.4380	1.8485		
$\frac{Y + Z}{2}$	C', G'	D', H'	B', F'	A', E'		
	1.7130	1.7515	1.3425	1.5580		

The computation of the "Position" and "Run" biases are shown in tables X and XI respectively. The columns are reordered to obtain a chain block design (see table I) based on the averages in the last two rows of table IX. Likewise, the rows are reordered to obtain a chain block

TABLE X.—POSITION EFFECTS

Position	I	III	II	IV	
j	1	2	3	4	
	A, E 1.8685 C', G' 1.7130	C, G 1.4380 B', F' 1.3425	B, F 1.9670 D', H' 1.7515	D, H 1.8485 A', E' 1.5580	
p_j	C', G'–C, G .2750	B', F'–B, F −.6245	D', H'–D, H −.0970	A', E'–A, E −.3105	$\bar{p} = -0.18925$
S	3	1	−1	−3	

$\hat{\gamma}_1 = \hat{\gamma}_{\mathrm{I}} = \frac{1}{8}[3.(.2750) + 1.(-.6245) - 1.(-.0970) - 3.(-.3105)] = +0.153625$
$\hat{\gamma}_2 = \hat{\gamma}_{\mathrm{III}} = +0.153625 - 0.2750 - 0.18925 = -0.310625$
$\hat{\gamma}_3 = \hat{\gamma}_{\mathrm{II}} = -0.310625 + 0.6245 - 0.18925 = +0.124625$
$\hat{\gamma}_4 = \hat{\gamma}_{\mathrm{IV}} = +0.124625 + 0.0970 - 0.18925 = +0.032375$

TABLE XI.—RUN EFFECTS

Run	W	Y	X	Z	
j	1	2	3	4	
	C, D 1.4675 A, B 1.8320	A', B' 1.3945 G', H' 1.5890	G, H 1.8190 E, F 2.0035	E', F' 1.5060 C', D' 1.8755	
p_j	A, B–A', B' .4375	G', H'–G, H −.2300	E, F–E', F' .4975	C', D'–C, D .4080	$\bar{p} = +0.27825$
S	3	1	−1	−3	

$\hat{\rho}_1 = \hat{\rho}_W = \frac{1}{8}[3.(.4375) + 1.(-.2300) - 1.(.4975) - 3.(.4080)] = -0.079875$
$\hat{\rho}_2 = \hat{\rho}_Y = -0.079875 - 0.4375 + 0.27825 = -0.239125$
$\hat{\rho}_3 = \hat{\rho}_X = -0.239125 + 0.2300 + 0.27825 = +0.269125$
$\hat{\rho}_4 = \hat{\rho}_Z = +0.269125 - 0.4975 + 0.27825 = +0.049875$

design. It should be pointed out that other reorderings could have been used for the rows or for the columns without altering the final results. The column biases are denoted by the letter γ, and the row biases by the letter ρ. The bias estimates $\hat{\gamma}_1$ and $\hat{\rho}_1$ are calculated according to equation (6), while the remaining three estimates of each set are obtained by the recursion formulas (7a) and (7b).

Table XII illustrates a convenient method for applying the position and run "corrections" to the observed values. The table to the right is obtained by adding to each entry in the left side table the corresponding row and column corrections. The corrections are the biases with their sign changed, rounded to three decimals to conform with the number of decimals in the original data.

TABLE XII.—CALCULATION OF CORRECTED VALUES

		I	II	III	IV	I	II	III	IV
Correction		−.154	−.125	+.311	−.032				
W	+.080	1.802	1.862	1.173	1.762	1.728	1.817	1.564	1.810
X	−.269	1.935	2.072	1.703	1.935	1.512	1.678	1.745	1.634
Y	+.239	1.610	1.568	1.267	1.522	1.695	1.682	1.817	1.729
Z	−.050	1.816	1.935	1.418	1.594	1.612	1.760	1.679	1.512

Table XIII lists the duplicate observations on each tire, both corrected and uncorrected. The table suggests that the design in this case was extremely effective in removing biases. This conclusion

TABLE XIII.—TREADWEAR OF TIRES

LOGARITHM OF TREADWEAR

Tire Symbol	Uncorrected		Corrected		Corrected Average
A	1.802	1.522	1.728	1.729	1.728
B	1.862	1.267	1.817	1.817	1.817
C	1.173	1.816	1.564	1.612	1.588
D	1.762	1.935	1.810	1.760	1.785
E	1.935	1.594	1.512	1.512	1.512
F	2.072	1.418	1.678	1.679	1.678
G	1.703	1.610	1.745	1.695	1.720
H	1.935	1.568	1.634	1.682	1.658

is confirmed by the analysis of variance shown in table XIV. Even though only 2 degrees of freedom are available for error, both the position and run effects are significant on better than the 5% level. It is interesting to convert back the position biases into antilogarithms, and note the large variation in rate of wear from one position to another.

TABLE XIV.—ANALYSIS OF VARIANCE OF TREADWEAR

Source	Degrees of Freedom	Sum of Squares	Mean Square
Total	15	.9680	—
Tires, uncorrected	7	.1844	.0265
Wheel Positions	3	.4281	.1427
Runs	3	.3486	.1162
Error	2	.0049	.0024

The sum of squares for "Wheel Positions" is obtained as follows:

$$2\{\tfrac{1}{2}[(.153625)\,(.2750 + .3105) + (-.310625)\,(-.6245 - .2750) + (.124625)\,(-.0970 + .6245) + (.032375)\,(-.3105 + .0970)]\} = .4281$$

The expression inside the braces is that given for S_3 in table VII, while the factor 2 is necessary because the data in table X are averages of two original observations. The calculation of the sum of squares for "Runs" is obtained in a similar way.

The hypothesis that all tires belong to the same population, from the viewpoint of rate of treadwear, can be tested by calculating the sum of squares of tires, corrected for position and run biases. First, sums of squares corresponding to rows and columns are calculated by the usual procedure for a two-way classification table, ignoring the tires. In the present case, one thus finds:

Sum of squares (uncorrected) for wheel positions (columns)

$$= 0.5150$$

Sum of squares (uncorrected) for runs (rows)

$$= 0.3592$$

The sum of squares for tires, corrected for positions and runs is then equal to "total $-$(error $+$ rows $+$ columns)":

$$= .9680 - .0049 - 0.5150 - 0.3592$$

$$= .0889$$

The corresponding mean square is

$$\frac{.0889}{7} = .0127$$

which is not significant on the 5% level in relation to the error mean square .0024.

On the other hand, it is known, from the data resulting from the entire test (16 runs) that there are real differences in rate of wear between some of these tires (which actually represent different brands). The small experiment here described failed to uncover these differences because of the low power of an F-test, having 2 degrees of freedom in the denominator. The complete test of 16 runs yields 194 degrees of freedom for error. The example is typical for situations in which systematic errors in the testing procedure (wheel positions, runs) are larger than most or all of the treatment differences. In these situations use of efficient statistical designs is not only helpful; it means the difference between a valid and a completely invalid experiment. It may be of interest to add that the effects of wheel positions, runs and tire differences computed from table IX are in good agreement with the corresponding values based on the entire test. The error term too is of the correct order of magnitude.

In an application like the one described here, attention must be given to the possibility of interactions which may invalidate the experiment. Three interactions must be considered: Tires $\times$ Wheel positions, Tires $\times$ Runs, and Wheel positions $\times$ Runs. Of these, the the last one can be eliminated if in the course of the test provisions are made not to disturb wheel alignment and other relevant features of the test vehicles, and to repeat the entire test in the case of an accident involving one or more of the vehicles. The interaction Tires $\times$ Runs has been found to be negligible provided that tread wear is determined by the weight method [2]. Finally, the interaction Tires $\times$ Wheel positions was also found to be small in most instances.

APPENDIX

1. *Derivation of Least Squares Solution for Two-Dimensional Chain Block.*

For purposes of simplicity of presentation, the notation used in the following departs somewhat from that used in the main body of the paper. No difficulty will result from this change in notation. Let y_α ($\alpha = 1, 2, \cdots, 4tq$) represent the observations. If the observation y_α occurs in row i and column j and corresponds to treatment k, we have:

$$y_\alpha = \mu + \rho_i + \gamma_j + \theta_k + \text{error} \tag{9}$$

where μ is a general mean, ρ_i a row effect, γ_j a column effect, and θ_k an effect of treatment.

Let $\hat{\mu}$, $\hat{\rho}_i$, $\hat{\gamma}_j$ and $\hat{\theta}_k$ be the least squares estimates of the corresponding parameters. To obtain these estimates it is convenient to consider the variables w_α , $x_{i\alpha}$, $u_{j\alpha}$ and $v_{k\alpha}$ such that

$$w_\alpha = 1 \text{ for all } \alpha$$

$$x_{i\alpha} = \begin{cases} 1 \text{ when } y_\alpha \text{ is in row } i \\ 0 \text{ otherwise} \end{cases}$$

$$u_{j\alpha} = \begin{cases} 1 \text{ when } y_\alpha \text{ is in column } j \\ 0 \text{ otherwise} \end{cases}$$

$$v_{k\alpha} = \begin{cases} 1 \text{ when } y_\alpha \text{ corresponds to treatment } k \\ 0 \text{ otherwise} \end{cases}$$

Let Y_α represent the regression value of y_α on w_α , all $x_{i\alpha}$, all $u_{j\alpha}$, and all $v_{k\alpha}$.

Then we have identically:

$$(10) \qquad Y_\alpha \equiv \hat{\mu} \cdot w_\alpha + \sum_i \hat{\rho}_i \cdot x_{i\alpha} + \sum_j \hat{\gamma}_j \cdot u_{j\alpha} + \sum_k \hat{\theta}_k \cdot v_{k\alpha}$$

The normal equations are:

$$(11a) \qquad \sum_\alpha y_\alpha = \sum_\alpha Y_\alpha$$

$$(11b) \qquad \sum_\alpha y_\alpha x_{i\alpha} = \sum_\alpha Y_\alpha x_{i\alpha} \qquad i = 1, 2, \cdots, 2q$$

$$(11c) \qquad \sum_\alpha y_\alpha u_{j\alpha} = \sum_\alpha Y_\alpha u_{j\alpha} \qquad j = 1, 2, \cdots, 2t$$

$$(11d) \qquad \sum_\alpha y_\alpha v_{k\alpha} = \sum_\alpha Y_\alpha v_{k\alpha} \qquad k = 1, 2, \cdots, 2tq$$

If $$\sum_i \hat{\rho}_i = \sum_j \hat{\gamma}_j = \sum_k \hat{\theta}_k = 0,$$

as is usually assumed,

(11a) becomes $$\sum_\alpha y_\alpha = 4tq\hat{\mu}$$

or

$$\hat{\mu} = \frac{\sum_{\alpha} y_\alpha}{4tq} = \bar{y} \tag{12}$$

Now, consider any one of the $2t$ equations (11c), say the equation corresponding to $j = j_0$. We have:

$$\sum_{\alpha} y_\alpha u_{j_0\alpha} = \sum_{\alpha} Y_\alpha u_{j_0\alpha} \tag{13}$$

Since $u_{j_0\alpha}$ is zero for all α for which y_α is not in column j_0, the summations in this equation extend only over all α for which y_α is in column j_0. Now, (13) can be written:

$$\begin{aligned}\sum_{\alpha} y_\alpha u_{j_0\alpha} &= \sum_{\alpha} [\bar{y} + \sum_{i} \hat{\rho}_i x_{i\alpha} + \sum_{j} \hat{\gamma}_j u_{j\alpha} + \sum_{k} \hat{\theta}_k v_{k\alpha}] u_{j_0\alpha} \\ &= \bar{y} \sum_{\alpha} u_{j_0\alpha} + \sum_{i} \hat{\rho}_i \sum_{\alpha} x_{i\alpha} u_{j_0\alpha} + \sum_{j} \hat{\gamma}_j \sum_{\alpha} u_{j\alpha} u_{j_0\alpha} \\ &\quad + \sum_{k} \hat{\theta}_k \sum_{\alpha} v_{k\alpha} u_{j_0\alpha}\end{aligned}$$

Carrying out the summations over all elements in column j_0, we obtain:

$$\begin{cases}\text{Sum of all observations in column } j_0 \\ \quad = 2q\bar{y} + \sum_{i} \hat{\rho}_i + 2q\hat{\gamma}_{j_0} + \text{sum of all } \hat{\theta}_k \text{ occurring in column } j_0 .\end{cases} \tag{14}$$

Since $\sum_i \hat{\rho}_i = 0$, this becomes, after division by $2q$.

$$\begin{cases}(\text{Average of all observations in column } j_0) - (\text{grand average}) \\ \quad = \hat{\gamma}_{j_0} + \text{average of all } \hat{\theta}_k \text{ occurring in column } j_0\end{cases} \tag{15}$$

There will be $2t$ such equations, corresponding to the $2t$ columns. Treating the equations of (11d) in a similar way, one obtains, for each k_0, an equation of the following type:

$$\begin{cases}(\text{Average of two observations on treatment } k_0) - (\text{grand average}) \\ \quad = \tfrac{1}{2}(\hat{\rho}_{i_1} + \hat{\rho}_{i_2}) + \tfrac{1}{2}(\hat{\gamma}_{j_1} + \hat{\gamma}_{j_2}) + \hat{\theta}_{k_0}\end{cases} \tag{16}$$

where i_1 and i_2 are the two rows in which k_0 occurs and j_1 and j_2 the two columns in which k_0 occurs.

Now, it is apparent from the design that if columns j_1 and j_2 have one treatment in common, they must have q treatments in common and that the $2q$ observations corresponding to these q treatments occupy

all $2q$ rows (each row once). Summing equations (16) for all these treatments, and dividing by q, one obtains:

$$(17)\quad \begin{cases} (\text{Average of all observations common to columns } j_1 \text{ and } j_2) \\ \qquad - (\text{grand average}) \\ = \dfrac{1}{2q}\sum_{i=1}^{2q}\hat{\rho}_i + \dfrac{1}{2}(\hat{\gamma}_{j_1} + \hat{\gamma}_{j_2}) \\ + (\text{average of all treatments, common to columns } j_1 \text{ and } j_2) \end{cases}$$

Since $\sum \hat{\rho}_i = 0$, we have:

$$(18)\quad \begin{cases} (\text{Average of all observations common to columns } j_1 \text{ and } j_2) \\ - (\text{grand average}) \\ = 1/2\,(\hat{\gamma}_{j_1} + \hat{\gamma}_{j_2}) + (\text{average of all treatments} \\ \text{common to columns } j_1 \text{ and } j_2). \end{cases}$$

There will be $2t$ equations of the type of equation (18). Let us now consider the semi-marginal averages of table III, as given in table IV. They can be considered as a new set of observations, $4t$ in number, and forming a design completely analogous to that of the original $4tq$ observations. In the new design, however, the value of q will be unity. The (true) parameters γ_j, denoting the column effects, are the same as in the original two-dimensional design, while the (true) row and treatment effects of the new design will be averages of sets of q corresponding row or treatment effects of the original design. By a line of reasoning exactly analogous to that leading to equations (15) and (18) we can obtain two corresponding sets of $2t$ equations each, say (15′) and (18′). It is readily seen, upon inspection of table III, that the first members of the equations (15′) and (18′) so formed will be numerically identical with those of the corresponding equation (15) or (18). The second members will contain new estimates of γ_j, say $\hat{\gamma}'_j$, and estimates for the averages of sets of q treatment effects. The matrix of these equations will be identical with that of (15) and (18). Consequently, in view of the numerical equality of the first members, the solutions will also be identical; that is, $\hat{\gamma}'_j = \hat{\gamma}_j$ for all j.

Thus, the least squares estimates of γ_j in the original design can be obtained by considering the simple chain-block of the sub-marginal totals, that is, by formulas similar to equation (6).

By interchanging rows and columns in table III it can be similarly shown that the least squares estimates for the row effects ρ_i are obtained from sub-marginal row averages (table V), which again form a simple chain block.

It remains to be shown that the least squares estimates for the treatment effects are obtained as indicated in section 2b. This is immediately evident from equation (16). Indeed, denoting by y_{α_1} and y_{α_2} the two observations for treatment k, (16) becomes equivalent to:

$$\bar{y} + \hat{\theta}_k = \tfrac{1}{2}\{[y_{\alpha_1} - (\hat{\rho}_{i_1} + \hat{\gamma}_{j_1})] + [y_{\alpha_2} - (\hat{\rho}_{i_2} + \hat{\gamma}_{j_2})]\}$$

Formulas (3) and (6), for the treatment and block effects of simple chain blocks are proved readily on the basis of the following consideration: In the simple chain block (table I), there are $2v$ equations of the type

$$y_\alpha = \mu + \rho_i + \gamma_j + \theta_k + \text{error} \tag{19}$$

The number of unknown parameters is composed of one parameter for the general mean μ, one for the row effects (since $\rho_1 + \rho_2 = 0$) $v - 1$ for the block effects (since $\sum \gamma = 0$) and $v - 1$ for the treatment effects (since $\sum \theta = 0$). Consequently, there being exactly as many equations as there are unknowns, the least squares solution is identical with the simple algebraic solution of set (19) (omitting the error term). It can be verified that relations (3) and (6) are indeed the solutions of equations (19), $\hat{\mu}$ being equal to $\bar{y}$, and $\hat{\rho}_1$ and $\hat{\rho}_2$ being the deviations of the row averages from $\bar{y}$.

2. *Variance of Treatment Differences*

The derivation of equation (8) proceeds along the following lines:

In accordance with equation (16), the difference between two treatment differences is of the form:

$$\hat{V}_1 - \hat{V}_2 = \hat{\theta}_k - \hat{\theta}_m = \tfrac{1}{2}[(A_k - A_m) + (R_k - R_m) + (C_k - C_m)]$$

where

A_k = sum of the two observations corresponding to treatment k

R_k = sum of the two row corrections for treatment k

C_k = sum of the two column corrections for treatment k and similar definitions for A_m , R_m , C_m .

Therefore:

$$\begin{aligned}\operatorname{Var}(\hat{V}_1 - \hat{V}_2) = {} & \tfrac{1}{4}[\operatorname{Var}(A_k - A_m) + \operatorname{Var}(R_k - R_m) + \operatorname{Var}(C_k - C_m)] \\ & + \tfrac{1}{2}[\operatorname{Cov}(A_k - A_m)(R_k - R_m) + \operatorname{Cov}(A_k - A_m)(C_k - C_m) \\ & + \operatorname{Cov}(R_k - R_m)(C_k - C_m)]\end{aligned}$$

We will prove that all the covariances vanish. The covariance of the R and C terms is zero because of the orthogonality of row and column corrections. This orthogonality is shown by the fact that the $2t$ equations (18), which entirely determine the column corrections, do not involve the row effects. The vanishing of the covariance of the A and C terms is seen from the following consideration:

C_k being the sum of block corrections, involves a sum of two expressions of the type (6), in which each term, according to (5) and (2b), involves the two observations of any treatment in the form of their difference. On the other hand, A_k and A_m involve treatments θ_k and θ_m only in the form of sums of duplicate observations. Since the covariance of the sum and the difference of two observations of equal precision is always zero, it follows that the A and C terms are uncorrelated. The same reasoning applies to the covariance of the A and R terms. If σ^2 denotes the variance of a single observation we have Var $(A_k - A_m) = 4\sigma^2$, and the two remaining terms necessary for the computation of Var $(\hat{V}_1 - \hat{V}_2)$ are readily computed on the basis of equations (5) and (6). Combining all results, equation (8) is obtained.

REFERENCES

[1] Fisher, R. A. and Yates, F., "Statistical Tables for Biological, Agricultural and Medical Research," New York, Hafner Publishing Company, Inc., 1948.

[2] Stiehler, R. D., Richey, G. G., and Mandel, J., Measurement of Treadwear of Commercial Tires, *Rubber Age, 73*, p. 201–208 (1953).

[3] Youden, W. J. and Connor, W. S., The Chain Block Design, *Biometrics, 9*, p. 127–140 (1953).

ANALYSIS FOR SOME PARTIALLY BALANCED INCOMPLETE BLOCK DESIGNS HAVING A MISSING BLOCK*

MARVIN ZELEN

National Bureau of Standards, Washington, D. C.

1. *Introduction and Summary*

The statistical design of experiments is becoming increasingly important in the physical sciences. This is especially true for those experiments where the block is the natural experimental unit. Thus in road testing automobile tires from different manufacturers, each automobile used in the tests can be regarded as a "natural" block; in conducting inter-laboratory tests, each laboratory can be a block; in fact almost every experiment in the physical sciences is characterized by the block being a "natural experimental unit".

However, as in all applications of experimental design, the experimenter will have to cope with unforeseen situations which may cause part of the experimental data to be missing. Since the block is the experimental unit, it is quite common in the physical sciences for an entire block to be lost. The problem of missing blocks has been discussed by other writers. The papers by Yates [1] and Yates and Hale [2] discuss the appropriate analysis for a Latin Square design. Anderson [3] has outlined the analysis of a split-plot design if a whole plot is lost. Cornish [4] gives the analysis for balanced incomplete block designs, having a missing block. This paper outlines the intra-block analysis (if a whole block is lost) for partially balanced incomplete block designs with two associate classes such that all treatments in the missing block are the same associates of each other.

2. *General Equations*

In all that follows the standard notation of experimental design will be used; v = number of treatments, r = number of times each treatment

*Presented before a joint session of The Biometric Society and the Institute of Mathematical Statistics in Washington, D. C. on May 1, 1953.

is replicated, b = number of blocks, k = number of experimental units per block. Then a partially balanced design with two associate classes is characterized by an experimental plan where no treatment occurs more than once in any block; and the treatments are arranged such that with respect to any particular treatment t the remaining treatments can be divided into two groups, each containing n_1 and n_2 treatments respectively so that treatment t occurs in λ_1 blocks with each of the treatments in the first group and in λ_2 blocks with each of the treatments in the second group. The treatments in each group are called the first and second associates of t respectively. Also if any two treatments are kth associates, the number of treatments common to the ith associates of one of the treatments and the jth associates of the other treatment is p_{ij}^k $(i, j, k = 1, 2)$, and is independent of the particular pair of treatments.

Now assume that the block containing treatments $t_1, t_2, \ldots, t_k$ is lost. Then if the adjusted yields for the ith treatment are defined by Q_i = (total yield for ith treatment) $-\frac{1}{k}$ (sum of the block yields in which the ith treatment occurs), the resulting normal equations can be written as

$$(2.1)\quad r(k-1)\hat{t}_i - \lambda_1 S_1(\hat{t}_i) - \lambda_2 S_2(\hat{t}_i) = kQ_i + k\hat{t}_i - \sum_{j=1}^{k} \hat{t}_j \text{ for } i \leq k$$

and

$$(2.2)\quad r(k-1)\hat{t}_i - \lambda_1 S_1(\hat{t}_i) - \lambda_2 S_2(\hat{t}_i) = kQ_i \quad \text{for} \quad i > k$$

where $S_j(\hat{t})$ = sum of the jth associates of t $(j = 1, 2)$.

From the theory of partially balanced incomplete block designs two additional equations defining a treatment estimate can be derived by summing the normal equations over the first and second associates of t_i. These can be written as

$$(2.3)\quad -\lambda_1 n_1 \hat{t}_i + S_1(\hat{t}_i)[r(k-1) - \lambda_1 p_{11}^1 - \lambda_2 p_{12}^1] + S_2(\hat{t}_1)[-\lambda_1 p_{11}^2 - \lambda_2 p_{12}^2] = kS_1(Q_i) + kS_1'(\hat{t}_i) - m_{1i} \sum_{j=1}^{k} \hat{t}_j ,$$

$$(2.4)\quad -\lambda_2 n_2 \hat{t}_i + S_1(\hat{t}_i)[-\lambda_1 p_{21}^1 - \lambda_2 p_{22}^1] + S_2(\hat{t}_i)[r(k-1) - \lambda_1 p_{21}^2 - \lambda_2 p_{22}^2] = kS_2(Q_i) + kS_2'(\hat{t}_i) - m_{2i} \sum_{j=1}^{k} \hat{t}_j ,$$

where $S_j(Q_i)$ $(j = 1, 2)$ is the sum of the adjusted treatment yields for those treatments which are jth associates of t_i , $S'_j(\hat{t}_i)$ is the sum of the jth associates of t_i occurring in the missing block and m_{ji} is the number of jth associates which treatment t_i has in the missing block.

Adding the restriction

$$\sum_{j=1}^{v} t_j = t_i + S_1(t_i) + S_2(t_i) = 0 \tag{2.5}$$

and following Bose and Shimamato [5], the solutions for the equations defining a treatment estimate occurring in the missing block can be written as

$$[r(k-1)-k]\hat{t}_i = kQ_i + c_1S_1(Q_i) + c_2S_2(Q_i) + c_1S'_1(\hat{t}_i) + c_2S'_2(\hat{t}_i) - \left[\frac{k + c_1m_{1i} + c_2m_{2i}}{k}\right] \sum_{j=1}^{k} \hat{t}_j \qquad \text{for} \qquad i \leq k \tag{2.6}$$

and the solutions for treaments not occurring in the missing block can be written as

$$r(k-1)\hat{t}_i = kQ_i + c_1S_1(Q_i) + c_2S_2(Q_i) + c_1S'_1(\hat{t}_i) + c_2S'_2(\hat{t}_i) - \left[\frac{c_1m_{1i} + c_2m_{2i}}{k}\right] \sum_{j=1}^{k} \hat{t}_j \qquad \text{for} \qquad i > k \tag{2.7}$$

where

$$c_1 = \frac{1}{k\Delta}\left[\lambda_1(rk - r + \lambda_2) + (\lambda_1 - \lambda_2)(\lambda_2 p_{12}^1 - \lambda_1 p_{12}^2)\right] \tag{2.8}$$

$$c_2 = \frac{1}{k\Delta}\left[\lambda_2(rk - r + \lambda_1) + (\lambda_1 - \lambda_2)(\lambda_2 p_{12}^1 - \lambda_1 p_{12}^2)\right] \tag{2.9}$$

and

$$\Delta = \frac{1}{k^2}\{(rk - r + \lambda_1)(rk - r + \lambda_2) + (\lambda_1 - \lambda_2)[r(k-1)(p_{12}^1 - p_{12}^2) + \lambda_2 p_{12}^1 - \lambda_1 p_{12}^2]\}. \tag{2.10}$$

The intra-block analysis of variance is shown in Table I where y_{ij} is the yield of the ith treatment in the jth block, B_j is the total yield of block j, G is the grand total of all the yields, and $N = bk$.

However, it is still not possible using equations (2.3), (2.4), (2.6), and (2.7) to solve explicitly for the treatment estimates. The quantities

$$S'_1(\hat{t}_i),\ S'_2(\hat{t}_i),\ \sum_{j=1}^{k} \hat{t}_j$$

TABLE I

INTRA-BLOCK ANALYSIS OF VARIANCE

Source of Variation	*Degrees of freedom*	*Sum of squares*	*Mean square*
Treatments (adjusted)	$v-1$	$S_t^2=\sum_{i=1}^{v}\hat{t}_iQ_i$	$s_t^2=\frac{S_t^2}{v-1}$
Blocks (unadjusted)	$b-2$	$S_b^2=\frac{1}{k}\sum_{i=1}^{b-1}B_i^2-\frac{G^2}{N-k}$	
Error	$N-b-v-k+2$	S_e^2 (by subtraction)	$s_e^2=\frac{S_e^2}{N-b-v-k+2}$
Total	$N-k-1$	$\sum_{i,j}y_{ij}^2-\frac{G^2}{N-k}$	

are unknown and will in general depend upon the particular design.

3. *Partially Balanced Designs with Two Associate Classes Such That All Treatments in the Missing Block Are the Same Associates Of Each Other.*

Assume that all treatments in the missing block are uth associates of each other. This special class of partially balanced designs includes all designs where one of the λ_i's is equal to zero and actually includes most of the partially balanced incomplete designs with two associate classes which are currently available.

Then if t_i is one of the treatments in the missing block ($i \leq k$)

$$S'_u(t_i) = \sum_{j=1}^{k} t_j - t_i \tag{3.1}$$

$$m_{ui} = k - 1 \tag{3.2}$$

and if w is the other type of associate ($u, w = 1, 2$)

$$S'_w(t_i) = 0 \tag{3.3}$$

$$m_{wi} = 0. \tag{3.4}$$

Then using the relations (3.1-3.4), equation (2.6) defining the treatment estimates in the missing block ($i \leq k$) can be simplified to

(3.5)
$$\hat{t}_i = \frac{1}{r(k-1)-k+c_u}\left\{kQ_i + c_1S_1(Q_i) + c_2S_2(Q_i) - \left(\frac{k-c_u}{k}\right)\sum_{j=1}^{k}\hat{t}_j\right\}$$

where c_u is defined by (2.8) or (2.9).

Summing equation (3.5) over all treatments in the missing block gives

$$\sum_{j=1}^{k}\hat{t}_j = \frac{1}{r(k-1)}\left\{k\sum_{j=1}^{k}Q_j + c_1\sum_{j=1}^{k}S_1(Q_j) + c_2\sum_{j=1}^{k}S_2(Q_j)\right\}. \tag{3.6}$$

Substituting (3.6) in (3.5) leads to the complete solution for estimates of the treatments in the missing block. Thus

$$\begin{aligned}\hat{t}_i = {} & \frac{1}{r(k-1)-k+c_u}\{kQ_i + c_1S_1(Q_i) + c_2S_2(Q_i)\} \\ & - \frac{[k-c_u]}{[r(k-1)-k+c_u][r(k-1)]k}\left\{k\sum_{j=1}^{k}Q_j + c_1\sum_{j=1}^{k}S_1(Q_j)\right. \\ & \left. + c_2\sum_{j=1}^{k}S_2(Q_j)\right\}.\end{aligned} \tag{3.7}$$

Once the estimates for treatments in the missing block have been solved explicitly using (3.7), it is possible to solve for the remaining treatment estimates using equation (2.7), where the quantities

$$S_1'(\hat{t}_i),\ S_2'(\hat{t}_i),\ \sum_{j=1}^{k}\hat{t}_j$$

are replaced by their numerical estimates.

In general the error terms for comparing the difference between two treatments will depend on whether both, one or none of the treatments occurs in the missing block; the number of different associates which each treatment has in the missing block; and the type of associate one treatment is in relation to the other.

If two treatments both occur in the missing block then

$$\text{var.}\,(\hat{t}_i - \hat{t}_j) = 2\left[\frac{k-c_u}{r(k-1)-k+c_u}\right]\sigma^2. \tag{3.8}$$

If two treatments are nth associates neither occurring in the missing block but such that both treatments have M_w common wth associates in the missing block and each has m_{wi}, m_{wi} wth associates in the missing

block, then

$$(3.9)\qquad \text{var.}\,(\hat{t}_i - \hat{t}_j) = \frac{\sigma^2}{B}\left\{2(k - c_n) + \frac{(c_1 - c_2)^2}{A}\left[(m_{wi} + m_{wj} - 2M_w) - \frac{(m_{wi} - m_{wj})^2}{k}\right]\right\}$$

where $A = r(k-1) - k + c_u$, $B = r(k-1)$.

If two treatments are nth associates such that (say) t_i occurs in the missing block and t_j does not occur in the missing block, then

$$(3.10)\qquad \text{var.}\,(\hat{t}_i - \hat{t}_j) = \sigma^2\left\{(k - c_n)\left[\frac{1}{A} + \frac{1}{B}\right] + [m_{1i}m_{2j}(c_1 - c_2)^2 + 2m_{wj}(c_w - c_u)(k - c_u) - (k - c_u)^2]\right\}$$

where A and B are defined as above.

4. *Illustrative Example*

In an experiment, X-ray diffraction patterns for tricalcium aluminate were recorded on different films, and in a number of instances the same X-ray reflection appeared on several of the films. If the intensity of each reflection is regarded as making a block and the different film responses as treatments, it is possible to determine if differential responses exist among the films. Table IIA summarizes the measurements (logarithm scale) and gives the experimental plan. The design is of the group divisible type except for the missing first block, and is catalogued as reference No. 3, p. 158 [5]. The association scheme for the design is

$$\begin{array}{ccc} a & b & c \\ d & e & f \\ g & h & i \\ j & k & l \end{array}$$

where treatments in the same column are first associates of each other and treatments in different columns are second associates of each other. The parameters of the design are

$$(4.1)\qquad v = 12,\quad b = 16,\quad r = 4,\quad k = 3,$$
$$\lambda_1 = 0,\quad \lambda_2 = 1,\quad p_{12}^1 = 0,\quad p_{12}^2 = 3,$$

from which it is possible to calculate $c_1 = 0$, $c_2 = \frac{1}{4}$.

TABLE IIA

Blocks	Observations and Experimental Plan						Block Totals
1	a	.....	b	.....	c	.....	.....
2	l	.3726	h	.6556	g	.2304	1.2316
3	i	.6402	j	.4622	e	.6716	1.7740
4	k	.3768	d	.3788	f	.5768	1.3324
5	a	.6556	l	.3098	k	.5186	1.4840
6	e	.4498	g	.4672	c	.4123	1.3293
7	f	.1746	j	−.0308	h	.0101	.1467
8	d	.5376	b	.5670	i	.5287	1.6333
9	a	.3990	e	.4609	f	.5086	1.3685
10	h	.2464	c	.2823	d	.2684	.7971
11	i	.2993	k	.1379	g	.1580	.5952
12	b	.1549	l	−.1320	j	−.0420	−.0191
13	a	.4622	h	.1858	i	.3788	1.0268
14	g	.3010	f	.5200	b	.4271	1.2481
15	j	−.0600	k	.0491	c	.0650	.0541
16	d	.5045	l	.2076	e	.4067	1.1188

Since all treatments in the missing block are second associates of each other, equation (3.7) giving the estimates for treatments occurring in the missing block, can be written as

$$(4.2) \qquad \hat{t} = \frac{4}{21}\left[3Q + \frac{1}{4}S_2(Q)\right] - \frac{11}{504}\left[3\sum_{j=1}^{3} Q_j + \frac{1}{4}\sum_{j=1}^{3} S_2(Q_j)\right]$$

where

$$\sum_{j=1}^{3} Q_j = Q_a + Q_b + Q_c .$$

Equation (2.7) defining the estimates for treatments not occurring in the missing block can be written as

$$(4.3) \qquad \hat{t} = \frac{1}{8}\left[3Q + \frac{1}{4}S_2(Q)\right] - \frac{1}{32}\left[\frac{m_2}{3}\left(\sum_{j=1}^{3}\hat{t}_j\right) - S_2'(t)\right]$$

where

$$\sum_{j=1}^{3}\hat{t}_j = \hat{a} + \hat{b} + \hat{c}.$$

Table IIB summarizes the treatment totals (T), the adjusted treatment yields (Q), the sum of the second associate Q's, $S_2(Q)$ and the treatment estimates calculated from equations (4.2) and (4.3). Table IIC summarizes the analysis of variance.

TABLE IIB

Treatments	T	Q	$S_2(Q)$	$\hat{t}$
a	1.5168	.2237	.3824	.1165
b	1.1490	.1949	−.2781	.0686
c	0.7596	.0328	−.1043	−.0158
d	1.6893	.0621	.3824	.0334
e	1.9890	.1255	−.2781	.0380
f	1.7800	.4147	−.1043	.1545
g	1.1296	−.3385	.3824	−.1169
h	1.0979	.0305	−.2781	.0024
i	1.8470	.1706	−.1043	.0630
j	0.3222	−.3297	.3824	−.1136
k	1.0824	−.0728	−.2781	−.0364
l	0.7580	−.5138	−.1043	−.1937

TABLE IIC—INTRA-BLOCK ANALYSIS OF VARIANCE

Source	Degrees of freedom	Sum of squares	Mean square
Treatments (adjusted)	11	.299846	.027259
Blocks (unadjusted)	14	1.528205	
Error	19	.128986	.006835
Total	44	1.957037	

This resulting design (with a missing block) will have five different error terms for comparing the differences between two treatment estimates. Table IID summarizes these different variances for typical treatment differences.

TABLE IID

Typical treatment differences	variance
$\hat{d}$–$\hat{e}$	.690 σ^2
$\hat{d}$–$\hat{g}$	.750 σ^2
$\hat{c}$–$\hat{e}$	.797 σ^2
$\hat{c}$–$\hat{d}$	.892 σ^2
$\hat{c}$–$\hat{b}$	1.045 σ^2

Acknowledgement: I would like to thank Dr. W. J. Youden for suggesting this problem, Dr. W. S. Connor, for many helpful suggestions, and Dr. F. Ordway and Mr. M. Grasso of the Portland Cement Fellowship at the National Bureau of Standards, for the use of their data in the illustrative example.

REFERENCES

[1] Yates, F. "Incomplete Latin Squares," *Journal of Agricultural Science, XXVI*, Pt. 2, (1936), 301–315.

[2] Yates, F. and Hale, R. W. "The Analysis of Latin Squares When Two or More Rows, Columns, or Treatments Are Missing," *Supplement Journal of the Royal Statistical Society, 6*, No. 1, (1939), 67–79.

[3] Anderson, R. L. "Missing-Plot Techniques", *Biometrics Bulletin, 2*, No. 3, (1946), 41–47.

[4] Cornish, E. A. "The Analysis of Quasi-Factorial Designs With Incomplete Data", *Journal of the Australian Institute of Agricultural Science, 6*, No. 1, (1940), 31–39.

[5] Bose, R. C. and Shimamato, T. "Classification and Analysis of Partially Balanced Incomplete Block Designs with Two Associate Classes", *Journal of the American Statistical Association, 47*, (1952), 151–184.

THE USE OF COVARIANCE TO CONTROL GRADIENTS IN EXPERIMENTS[1]

W. T. Federer and C. S. Schlottfeldt[2]

Cornell University

INTRODUCTION

The purposes of this paper are to illustrate the use of covariance to control gradients in experimental material with an actual example, to discuss the use of covariance instead of stratification to control variation, and to indicate some possible applications of the procedure.

THE EXAMPLE AND ANALYSIS

In the spring of 1951 an experiment was devised to determine whether the exposure of tobacco seeds to different dosages of cathode rays would affect the growth of the resulting plants. The seeds were from a strain of tobacco which had been under controlled pollination since 1909, and, hence, the material used in the experiment was highly uniform with respect to its genetical background. The seven different treatments (the different dosages of cathode rays) were laid out in a randomized complete block experiment with eight replicates. The plot size was 2 rows by 10 plants with 3 feet between rows and 1.5 feet between the plants. The following measurements were made on each plant:

(i) Plant height on 7/13/51 in cm. equals 1st plant height.
(ii) Plant height on 8/14/51 in in. equals 2nd plant height.
(iii) Length of longest leaf on 7/13/51 in cm. equals 1st leaf length.
(iv) Length of longest leaf on 8/14/51 in in. equals 2nd leaf length.
(v) Width of widest leaf on 7/13/51 in cm.

The available information indicated that the experimental area was uniform within the replicates but not between replicates. Shortly after the plants were transplanted to the field it became apparent that an environmental gradient existed from the center of the replicates outward. (see figure 1). The soil fertility decreases from replicate one

[1]Paper no. 301 of the Department of Plant Breeding and no. 13 of the Biometrics Unit. The authors are indebted to Dr. H. H. Smith for the use of these data.

[2]Now at Universidade Rural, Viçosa, Minas Gerais, Brazil.

through replicate eight. The gradient across the treatment plots within the replicate is curvilinear with the bottom part of the curve lying in the center of the replicates. The gradients in the experimental area had a marked effect upon the characters measured.

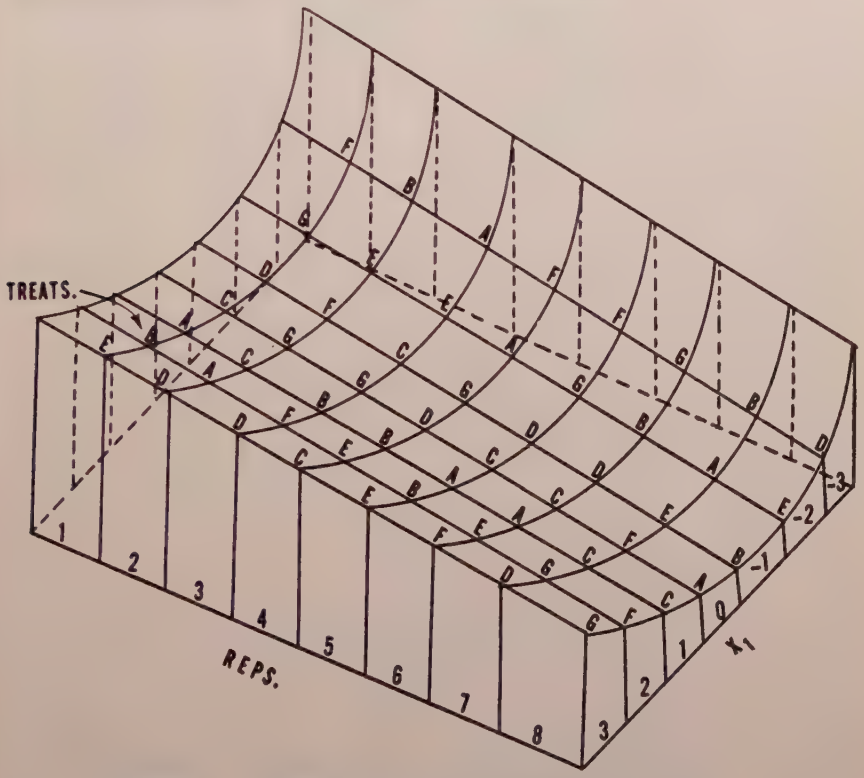

FIGURE I. DIAGRAMMATIC REPRESENTATION OF GRADIENTS AND LOCATION OF TREATMENTS (A, B, ...G) IN EACH REPLICATE.

The initial data for 1st plant height in the original field arrangement are given in table I, and the analysis of variance on these data is presented in table II. The coefficient of variation, $56\sqrt{30228.2}/56698.6 = 17.2\%$, is quite high for material of this type. In order to control some of the variability due to the gradient across the treatment plots within the replicate, a covariance analysis on position of the treatment within a replicate was used. The columns in figure 1 were numbered

TABLE I

PLANT HEIGHT, FIRST MEASUREMENT (TOTALS FOR 20 PLANTS IN CM.)

Covariate		Replicate number								Column totals
X	X^2	1	2	3	4	5	6	7	8	
−3	9	*F* 1299.2	*B* 1369.2	*A* 1169.5	*F* 1219.1	*F* 1120.0	*G* 1031.5	*B* 1076.4	*D* 1099.6	9384.5
−2	4	*G* 875.9	*E* 844.2	*E* 975.8	*A* 971.7	*G* 827.0	*B* 846.5	*A* 917.9	*E* 947.4	7206.4
−1	1	*D* 960.7	*F* 968.7	*C* 873.4	*G* 607.6	*D* 671.9	*D* 667.8	*E* 627.6	*B* 787.1	6164.8
0	0	*C* 1004.0	*G* 975.5	*G* 797.8	*D* 1000.0	*C* 972.2	*C* 853.6	*F* 776.4	*A* 898.3	7277.8
+1	1	*A* 1173.2	*C* 1322.4	*B* 1069.7	*B* 1343.3	*A* 1083.7	*A* 1087.1	*C* 960.4	*C* 1174.9	9214.7
+2	4	*B* 1031.9	*A* 1172.6	*F* 1093.3	*E* 999.4	*B* 1146.9	*E* 990.2	*G* 852.4	*F* 1003.3	8290.0
+3	9	*E* 1421.1	*D* 1418.9	*D* 1169.6	*C* 1181.3	*E* 993.8	*F* 1021.9	*D* 1006.2	*G* 947.6	9160.4
Total		7766.0	8071.5	7149.1	7322.4	6815.5	6498.6	6217.3	6858.2	56698.6

TABLE II

ANALYSIS OF VARIANCE

Source of variation	Degrees of freedom	Mean square	F
Total	55	35,123.0	
Replicates	7	55,473.6	1.84
Treatments	6	45,645.9	1.51
Error	42	30,228.2	

-3, -2, -1, 0, 1, 2, and 3 instead of 1, 2, 3, 4, 5, 6, and 7 in order to simplify the analysis of covariance. The use of the former sequence of numbers adds considerably to the simplification of the calculations since the sequence adds to zero, thus eliminating the correction terms, and the relationship between these numbers and their squares is zero. The symbol X_1 is used to denote the numbers in the sequence, and the symbol X_2 is used to denote the squares of the numbers.

For illustrative purposes the analysis of covariance on the linear trend across the replicates is given in table III. Due to the nature of

TABLE III

LINEAR COVARIANCE ANALYSIS

Source of variation	Sum of products				Errors of estimate		
	D.F.	Σx^2	Σxy	Σy^2	D.F.	Sum of squares	Mean square
Total	55	224.00	4,544.800	1,931,766.7			
Reps.	7	0	0	388,314.9			
Treats.	6	9.25	−14.875	273,875.4			
Error	42	214.75	4,559.675	1,269,586.4	41	1,172,773.2	28,604.22
$E + T$	48	224.00	4,544.800	1,543,461.8	47	1,451,251.1	
Treatments adjusted for linear regression					6	278,477.9	46,412.98

the curvilinear relationship, little reduction in the error mean square is obtained for the linear covariance on trend across the plots. Upon fitting a curvilinear covariance of second degree (table IV) a con-

TABLE IV

QUADRATIC COVARIANCE ANALYSIS*

Source of variation	Sum of products						
	D.F.	Σx_1^2	Σx_1x_2	Σx_2^2	Σx_1y	Σx_2y	Σy^2
Total	55	224.00	0.0	672.00	4544.800	17,474.80	1,931,776.62
Reps.	7	0	0	0	0	0	388,314.90
Treats.	6	9.25	−9.5	86.75	−14.875	−431.75	273,875.44
Error	42	214.75	9.5	585.25	4559.675	17,906.55	1,269,586.28
$E + T$	48	224.00	0.0	672.00	4544.800	17,474.80	1,543,461.71

Source of variation	Errors of estimate			
	D.F.	Sum of squares	Mean square	F ratio
Error	40	687,793.47	17,194.8	
Error + Treats.	46	996,833.32	.....	
Treats. adj. for regression	6	309,039.85	51,506.6	3.00

*x_1 is used to refer to the covariate and x_2 to the squares of the covariate.

siderable reduction in the error mean square is obtained. In fact, the error mean square is little more than half that obtained in table II.

Sir Ronald A. Fisher describes a covariance analysis in which the linear gradient within the replicates is taken into account.[1] In addition, he presents a method for determining the gain in information from the various procedures. For our example, let us suppose that a standard error of the mean equal to five per cent of the mean is our criterion of precision for this experiment. In order to obtain the assumed error variance, square five per cent of the mean $(.05 \times 1012.5)^2 = 2562.890625$, and multiply the result by the number of replicates; thus, $8(2562.890625) = 20503.125$. The amount of information is equal to the reciprocal of the error variance, and the efficiency of two procedures is the ratio of the amounts of information. The efficiency of the randomized block design without covariance relative to an experiment with the assumed error variance is the ratio of the two amounts of information, $1/30228.2 \div 1/20503.125 = 20503.125/30228.2 = .6783$ unit of information. Since the error mean square, 30228.2, is estimated with 42 degrees of freedom, the fractional loss in information due to

[1]Statistical Methods for Research Workers, section 48, 10th edition.

estimating this mean square is 2/(error d.f. + 3) = 2/45. Therefore, the total unit of information is (1 − 2/45)(.6783) = .65, which is the first value given in table V. The remaining values are computed in a

TABLE V

GAINS IN UNITS OF INFORMATION

Character		Type of analysis		
		Variance	Linear covariance	Quadratic covariance
Plant height	1st	0.65	0.68	1.14
	2nd	3.68	3.96	5.36
Leaf length	1st	3.53	3.55	8.40
	2nd	3.28	3.24	5.12
Leaf width	1st	2.33	2.33	4.97
Average		2.69	2.75	5.00

similar manner for the other analyses and for other characters. Units of information computed in this manner have the same invariant properties as the coefficient of variation.

Although the hypothesis of equality of the seven treatment means is not tenable for this particular selection of treatments (one of the treatments represents a control; i.e., the seeds were not exposed to cathode rays), it is interesting to note the effect of the covariance analyses on the F ratios in table VI. None of the F values are close to the tabulated

TABLE VI

F VALUES—RATIO OF TREATMENT TO ERROR MEAN SQUARES

($F_{05}(6, 40 \text{ df}) = 2.34$; $F_{01}(6, 40 \text{ df}) = 3.26$)

Character		Type of analysis		
		Variance	Linear covariance	Curvilinear covariance
Plant height	1st	1.51	1.62	3.00
	2nd	1.04	1.29	2.07
Leaf length	1st	0.76	0.72	2.90
	2nd	0.30	0.33	0.76
Leaf width	1st	0.79	0.87	2.62

five per cent value for F when account is not taken of the curvilinear gradient within the replicate. With a second degree curvilinear covariance analysis four of the five F values are near or beyond the five per cent value for F. The remaining F value is much nearer unity than it was before accounting for the gradient within the replicates.

The adjusted treatment means are obtained from the following formula:[1] $\bar{y}_i' = \bar{y}_i$ (adjusted) $= \bar{y}_i$ (unadjusted) $- b_{y1.2}(\bar{x}_{1i} - \bar{x}_1) - b_{y2.1}(\bar{x}_{2i} - \bar{x}_2)$, where

$$b_{y1.2} = \frac{E_{22}E_{y1} - E_{12}E_{y2}}{E_{11}E_{22} - E_{12}^2}$$

$$= \frac{585.25(4559.675) - 9.5(17906.55)}{214.75(585.25) - 9.5(9.5)} = 0.198933,$$

$$b_{y2.1} = \frac{E_{11}E_{y2} - E_{12}E_{y1}}{E_{11}E_{22} - E_{12}^2}$$

$$= \frac{214.75(17906.55) - 9.5(4559.675)}{214.75(585.25) - 9.5(9.5)} = 30.27350,$$

$\bar{x}_1 = 0$, $\bar{x}_2 = 8(9 + 4 + 1 + 0 + 1 + 4 + 9)/56$, and the E_{11}, E_{22}, E_{12}, E_{y1}, and E_{y2} are the various sums of squares and cross products in the error line of the analysis of variance table (table IV). The adjustments, unadjusted totals, and the adjusted means are given in table VII.

The standard error of the difference between two adjusted treatment means, say $\bar{y}_i' - \bar{y}_j'$, is

$$\left(\text{Error mean square} \cdot \left\{\frac{2}{r} + \frac{E_{22}(\bar{x}_{1i} - \bar{x}_{1j})^2 - 2E_{12}(\bar{x}_{1i} - \bar{x}_{1j})(\bar{x}_{2i} - \bar{x}_{2j}) + E_{11}(\bar{x}_{2i} - \bar{x}_{2j})^2}{E_{11}E_{22} - E_{12}^2}\right\}\right)^{1/2}$$

$$= \left(17194.8 \cdot \left\{\frac{2}{8} + \frac{585.25(\bar{x}_{1i} - \bar{x}_{1j})^2 - 2(9.5)(\bar{x}_{1i} - \bar{x}_{1j})(\bar{x}_{2i} - \bar{x}_{2j}) + 214.75(\bar{x}_{2i} - \bar{x}_{2j})^2}{214.75(585.25) - 9.5(9.5)}\right\}\right)^{1/2}$$

DISCUSSION

The use of covariance to control the gradient across the treatment plots approximately doubled the amount of information obtained from

[1] J. Wishart, Tests of significance in analysis of covariance, J. Roy. Stat. Soc., Suppl. 3:79–82, 1936.

TABLE VII

UNADJUSTED AND ADJUSTED TREATMENT TOTALS AND MEANS

Treatment	Total			Adjustments for total		Adjusted	
	$X_{1i.}$	$X_{2i.}$	$Y_{i.}$	$b_{y1.2}(X_{1i.} - 0)$	$b_{y2.1}(X_{2i.} - 32)$	Total	Mean
A	−2	24	8474.0	−0.398	−242.188	8231.414	1028.93
B	−3	33	8671.0	−0.597	30.273	8700.676	1087.58
C	5	13	8342.2	0.995	−575.196	7767.999	971.00
D	3	39	7994.7	0.597	211.914	8207.211	1025.90
E	3	39	7799.5	0.597	211.914	8012.011	1001.50
F	−3	45	8501.9	−0.597	393.556	8894.859	1111.86
G	−3	31	6915.3	−0.597	− 30.273	6884.430	860.55
Total	0	224	56698.6	0.000	0.000	56698.600	.....
Mean	0	32	1012.5	.....	.....	.....	1012.5

the experiment. (table V). In order to attain the same precision about twice as many replicates would be required when the effect of the gradient is not removed by covariance. The rather large gains attained in this experiment may also be found in certain other types of experiment. For example, a linear gradient or a curvilinear gradient of second degree (either convex or concave) may exist in (i) greenhouse experiments where the source of heat is located on the sides of the house, (ii) field experiments located in areas containing drainage tiles, (iii) field experiments containing a depression in the center of the replicates, (iv) orchard and vineyard experiments on undulating topography, (v) animal experiments with the animals located at varying distances from the source of heat, (vi) experiments in which the yields are affected by slowly migrating insects entering the area from one side, etc.

The latin square effectively controls the sort of variation described above. In some cases a more effective control may be obtained with the latin square than with covariance while in other cases the use of covariance may prove more effective. An example of the latter situation is provided in experiments where the material is divided into size groups. Covariance on actual size will usually be more effective than using rough groupings of size for the rows or colums of a latin square. In addition, the utilization of covariance does not use up as many degrees of freedom as does stratification into rows or columns. For the present example only two degrees of freedom are required for the covariance analysis while six degrees of freedom would be required with a 7×7 latin square design.

The decision to use covariance to control gradients after the experimental results have been studied invalidates the use of tabulated probability values for the standard tests of significance. If the decision to use covariance to control gradients is made prior to conducting the experiments the tabulated probability values may be used for comparison with observed values in the standard tests of significance. Although the general decrease in plant vigor in the center of the replicates was noted shortly after transplanting, it is doubted that this observation unduly affects the use of tabulated probability levels for standard tests of significance for these data. If information concerning the gradient had been available prior to transplanting, a latin square design would have been chosen instead of the randomized complete block design.

DESIGN AND ANALYSIS OF SOIL INSECTICIDE FIELD EXPERIMENTS

D. VAN DER REYDEN

Tobacco Research Board, Salisbury, Southern Rhodesia, Africa

1. *Summary.*

A method of adjustment based on the linear hypothesis of the analysis of variance of a set of control plots to adjust associated treated plots for uneven distribution of soil insects is described, and applied to a soil insecticide experiment.

2. *Introduction*

In designing ordinary field experiments the devices of replication and randomization ensure an unbiased and statistically precise evaluation of effects. Under the null hypothesis, the probability of obtaining a significant result depends on the magnitude of the inherent variation of the variable tested, the number of replications, and the number of degrees of freedom the error variance is based upon. The smaller the, usually unknown, interaction between treatment and soil variations the smaller the estimate of the true error variance. Accordingly the experimenter attempts to select as uniform a piece of land as possible and prefers to use designs with small block size so as to reduce unavoidable soil variations. Under these conditions the assumptions of the experimental model applied are usually approximately satisfied, viz. that the various effects and error are additive, that the errors are the same from one plot to another, non-correlated, and normally distributed.

In field experiments where soil insecticides and related treatments are to be compared, the determining factors are the supply and distribution of the soil insects. The experimenter has to ensure, firstly, that there is an abundant supply of insects in the experimental field, a requirement that may severely limit the size of the experiment and consequently the number of replications. Even if this requirement can be met, the distribution of soil insects may be so irregular that the assumptions on which the experimental model rests may no longer be realistic.

It will be shown that by attaching a control plot to *each* treated plot the data from the control plots can be utilized to adjust the data from

the treated plots, so that an analysis of variance of the adjusted treated plots becomes valid.

3. *The Use of Control Plots*

Since it is difficult in field experiments of this nature to measure the effects of insecticides directly, some indirect measurement based on damage or loss of plants of a test crop is usually employed.

Let W be the measurement of activity of the soil insects in an individual control plot, and U that of the associated individual treatment plot. If the soil insects, as measured by their responses in the control plots, are non-randomly distributed over the experiment, estimates of the true measurements for a set of individual plots—consisting of a control plot and a treated plot may be written

$$\hat{X} = X_p + \epsilon_x + d_p \tag{1}$$

where the symbol X can be read either as W or U, $\hat{X}$ denote the value that would have been obtained had the soil insects been randomly distributed, X_p that part of the measurement associated with the experimental design, ϵ_x that part independent of the design, and d_p the magnitude of the correction.

If the experiment is properly randomized, block size small, and the linear hypothesis satisfied, it can be assumed that ϵ_w and ϵ_u are chance variables and estimates of a common random variable ϵ. This being the case it follows that the average of ϵ_w and ϵ_u, viz. $\bar{\epsilon}$, is a better estimate of the random variable than any single one. Furthermore a valid estimate of d_p can be obtained by applying the method of least squares to the linear hypothesis determined by the experimental design.

To prepare the way for the practical example used later in this paper, assume that the experimental design is a simple rectangular lattice. By putting the subscript p equal to eij, X_p in equation (1) can be written

$$X_{eij} = m(X) + r_e(X) + b_{ei}(X) + t_j(X) + \epsilon_{eij}(X) \tag{2}$$

where $e = 1, 2$ replicates; $i = 1, \cdots, k$ incomplete blocks; $j = 1, \cdots, k(k-1)$ treatments, and m, r, b and t are the constants for general mean, replicate, incomplete block, and treatment respectively.

The data from both the control and treated plots viz. W and U are subjected to the ordinary analysis of variance test to determine the significance, or otherwise, of these constants. The control plots are analysed as if they were treated in the same way as their associated treated plots. If the soil insects were randomly distributed, "treatments" effect for the control plots would be insignificant in the analysis

of variance. The significance of "treatments" for W thus affords a criterion for the randomness of distribution of the soil insects in the experimental field and, at the same time, a method for estimating d in (1). If $t_i(W)$ is insignificant no further attention need be given to the data from the control plots. If it is significant $t_i(W)$ is an estimate of the correction factor d and for practical purposes can be used as such.

Proceeding with the adjustment calculate the constants in (2) using the formulae supplied in the next section. After correction the following formal situation will be obtained

$$\begin{aligned} \hat{U}_{esij} &= m(U) + r_e(U) + b_{esi}(U) + t_j(U) + (d_j) + \bar{\epsilon}_{esij} \\ \hat{W}_{esij} &= m(W) + r_e(W) + b_{esi}(W) + \bar{\epsilon}_{esij} \end{aligned} \tag{3}$$

$\hat{W}$, after elimination of replicate and block effects, will now be the random variable the experimenter required in the first place, and an analysis of variance of $\hat{U}$ will accurately reflect the influence of the various treatments on the soil insects.

4. *Estimation of Experimental parameters*

The calculations required will only be indicated here using the 6×7 simple rectangular lattice as example. The calculations are set out according to the standard scheme in Robinson and Watson (1949) or Harshbarger (1947, 1949). Following Harshbarger, typical formulae for obtaining the estimates of the parameters on an intra-block basis are

$$\begin{aligned} 84m &= G \\ 84r_e &= 2R_e - G \\ 35b_{x1} &= 6(B_{x1} - T_{y1}) - (B_{y2} - T_{x2}) - (R_1 - R_2) \\ 35b_{y1} &= 6(B_{y1} - T_{x1}) - (B_{x7} - T_{y7}) + (R_1 - R_2) \\ 2t_1 &= T_1 - 2m - b_{x1} - b_{y1} \end{aligned} \tag{4}$$

where G represents the total of all observations, R_e represents the total of observations in replicate e, B_{x1} is the marginal total for incomplete block 1 in the X-replicate, T_{y1} is the marginal total for the symbols in block $X1$ but summed over the Y-replicate, and T_1 is the total of treatment 1 over the two replicates. The ϵ's are obtained by subtraction according to (2).

5. *Practical Example*

In the 1952/53 season an experiment was carried out at the Trelawney Tobacco Research Station by Miles (1953) to compare the efficacy

of various chemical and cultural control measures on the main soil insects that attack tobacco in Southern Rhodesia. These are whitegrubs and false wireworms, the larvae of *Rutelids* and *Melolonthids* and of *Tenebrionids* respectively.

To ensure an ample supply of larvae, heavily manured land was chosen. This choice naturally limited the size of the experiment. Apart from two chemicals (Chlordane and gamma BHC), each at four rates, applied at four different times, there were 10 cultural treatments consisting of five different times of planting tobacco with and without one level of one chemical, making up 42 treatment-combinations in all. The treatments were randomized according to a 6×7 simple rectangular lattice, thus ensuring small block size and equal precision of effects.

TABLE 1. OBSERVED AND CORRECTED STAND LOSS PERCENTAGES

Time of Application	Chlordane (lbs/acre) 0.75	1.5	2.25	3.00	gamma BHC (lbs/acre) 0.1	0.2	0.3	0.4	Time of Planting	BHC None	0.4	Explanation
	(16)	(11)	(25)	(6)	(7)	(31)	(12)	(1)		(9)	(5)	Symbol
	34	49	45	49	48	42	49	45		46	31	Control
10/11/52	28	25	8	22	26	22	30	25	10/11/52	12	27	Treated
	19	27	16	34	32	15	35	34		17	17	Corrected
	(30)	(14)	(24)	(37)	(21)	(8)	(23)	(18)		(34)	(28)	Symbol
	35	44	32	46	22	39	36	49		53	39	Control
24/11/52	35	26	14	25	19	30	36	41	24/11/52	23	17	Treated
	31	28	7	29	7	33	27	48		29	20	Corrected
	(4)	(22)	(36)	(29)	(13)	(20)	(38)	(40)		(3)	(10)	Symbol
	42	40	66	33	41	46	38	43		30	57	Control
8/12/52	32	18	19	24	38	28	24	23	8/12/52	23	41	Treated
	36	18	35	15	38	36	20	26		19	53	Corrected
	(33)	(27)	(35)	(41)	(32)	(17)	(19)	(42)		(26)	(2)	Symbol
	41	43	47	33	41	40	41	33		30	31	Control
22/12/52	25	22	32	12	26	24	23	32	22/12/52	33	41	Treated
	21	32	28	8	17	18	26	18		25	40	Corrected
	Except where indicated all plots planted on 10/11/52									(15)	(39)	Symbol
										39	42	Control
	L.S.D. Corrected Stand losses: 13% ($P = .05$) 18% ($P = .01$)								5/1/53	18	27	Treated
										18	31	Corrected

Stand loss of tobacco was taken as measurement of larvae activity. To measure the effects of possible uneven distribution of larvae and associated factors, every treatment plot consisting of two rows of 27 tobacco plants each was flanked by a control row with the same number of plants. These control rows together were considered as an individual control plot for the treatment they enclosed, and thus an arrangement of control plots corresponding with the arrangement of the treated plots was obtained as indicated in figure 1. This arrangement was chosen because prior experimental evidence showed that there was little, if any, movement of the grubs. It could only be conluded that the unequal distribution was due to preferences of the female in her choice of oviposition sites (Miles, 1953).

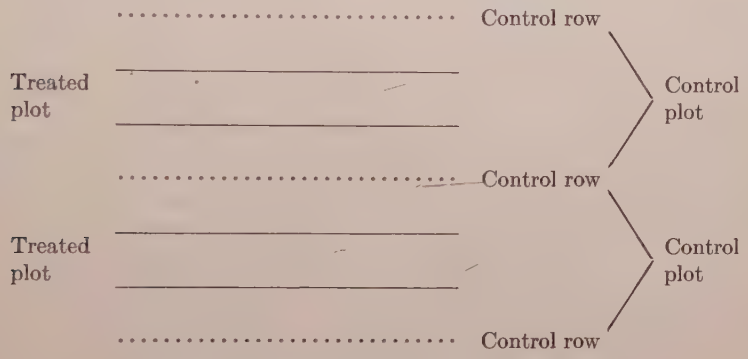

FIGURE 1. DIAGRAM OF PLOT ARRANGEMENT

Table 1 summarizes the treatments and their measurements. The bracketed figures are the symbols for the various treatment-combinations. The first row of figures below the symbols are the totals of the stand losses in the control plots, while the second row of figures are the observed stand losses of the treated plots. Since stand losses were scored out of a maximum of 50 plants per plot (end plants in each row being discarded) these figures being the totals of two replicates are automatically the percentage stand losses.

Analysing stand losses for both the control and treated plots in the usual way (Cochran & Cox, 1950), the analysis of variance is presented in Table II. In the case of the treated plots not a single effect attained significance, while both blocks and "treatments" showed significance for the control plots. This contradiction demonstrates the non-random distribution of larvae and associated factors causing stand loss. Equation (1) is therefore operative, and no comparisons can be made between the results of the treated plots until these are corrected for the bias.

TABLE II. ANALYSIS OF VARIANCE OF STAND LOSSES

Source	D.F.	Variances			Variance Ratios		
		Control	Treated	Corrected	Control	Treated	Corrected
Replicates	1	46.0	66.0	61.0	2.84	1.84	6.48*
Blocks (adj.)	12	41.8	30.5	32.3	2.57*	·	3.43**
Chemicals (C)	1	9.0	99.0	46.0	·	2.75	4.88*
Rates Chlordane (R)	3	35.0	51.0	25.0	2.16	1.42	2.66
Rates BHC (R')	3	8.3	6.3	22.0	·	·	2.34
Application Time (T)	3	42.0	18.3	47.6	2.59	·	5.06**
CT	3	4.7	1.7	6.0	·	·	·
RT	9	44.7	26.6	75.8	2.75*	·	8.06**
$R'T$	9	24.6	25.1	69.2	1.51	·	7.35**
Factorial Treatments (F)	31	29.1	25.7	53.3	1.79	·	5.66**
BHC Plantings (P)	4	50.8	30.0	13.0	3.13	·	1.38
Control Plantings (P')	4	57.0	53.5	109.5	3.51	1.49	11.64**
F vs P vs P'	2	7.0	50.0	81.0	·	1.39	8.61**
All Treatments	41	32.8	30.0	56.2	2.02*	·	5.97**
Error	29	16.2	36.0	9.4			
General Mean					20.6	12.8	12.8
Coeft. of Variation (%)					19.6	43.8	24.0

*significant at $P = .05$
**significant at $P = .01$

The constants are, therefore, calculated for both the treated and the control plots according to (4) and the corrected stand loss, $\hat{U}$, constructed according to (3). Note that $\bar{\epsilon}$ is the average of the ϵ's of the treated and the control plots. A summary of the corrected stand losses is given in the third row of Table I, and their analysis is given in Table II.

High significance is now obtained for treatment and other effects of the corrected treated plots. It is interesting to observe the reduction of the Coefficient of Variation from 44% to 24%.

6. *Comments*

As was stated before, the layout in figure 1 was chosen because prior evidence showed that there was no appreciable movement of the soil insects. This layout has the advantage that many treatments can be

compared in the experiment, the size of which is limited by insect supply. If in ect movement is suspected, however, it would be necessary to incorporate guard rows. A possible disadvantage of the layout, and this was not realized at the time of planning, is the danger of introducing an element of correlation in the adjusted treated plots by using the control row twice over as indicated in Figure 1. This difficulty could have been overcome in this experiment by using two control rows per two treatment rows.

At first sight it might be thought that adjustments could be effected by means of an analysis of covariance. A moment's reflection would show, however, that covariance techniques are meaningless in this problem, since the more effective a treatment the less the correlation between treated and control plots. The correlation coefficient obtained in this experiment was 0.2079, while the value required at the 5% level is 0.3500.

The uses of this method of adjustment are not restricted to soil insecticide experiments. In effect, the one control to one treated plot approach could be interpreted as a uniformity or calibration trial executed concurrently with the real trial. This concurrency would be especially useful with crops which should not normally be planted continuously on the same site.

In general the large reduction of variation and the higher precision thus obtained should compensate for the labour of recording and analysing the additional measurements.

7. *Acknowledgement*

This paper is published with the permission of Dr. F. A. Stinson, Director of the Tobacco Research Board of Rhodesia, to whom I am grateful for encouraging this type of research.

REFERENCES

Cochran, W. G. and G. M. Cox. *Experimental Designs,* N. Y. John Wiley 1950.

Harshbarger, Boyd. Rectangular Lattices. *Virginia Agr. Expt. Sta. Mem. 1* 1947.

———. Triple Rectangular Lattices. *Biometrics 5*(1): 1, 1949.

Miles, P. W. Problems in Soil Insecticide Research in Rhodesia. *S. Afr. J. Sci.* 1953.

Robinson, H. F. and G. S. Watson. An analysis of simple and triple rectangular lattice designs. *North Carolina Agr. Expt. Sta. Tech. Bul.* No. 88 1949.

QUERIES

GEORGE W. SNEDECOR, EDITOR

108 QUERY: Our geneticists like to have a check plot (X), carrying the standard variety of the area in which they are testing seedlings, adjacent to every seedling plot, since this greatly assists them in making their frequent plot-to-plot seedling gradings (throughout the 20-24 months cropping period) against the standard variety which the seedling must be able to beat in many plant characteristics as well as yields. Thus, a typical block plan for testing 5 seedlings (A, B, C, D, E) and their check (X) might be like this:

C	*E*	(*X*)	*B*	*A*	(*X*)
(*X*)	*D*	*A*	(*X*)	*E*	
B	(*X*)	*C*	*D*	(*X*)	

etc. with 8 plots in each Block

Now I am not sure how to handle the "missing-plot" formula when the missing data happen to be from *one* of the *three* X plots in one of the Blocks.

Herewith is a specific case.

Blocks	Standard Variety			Seedlings					Block Totals
	X			*A*	*B*	*C*	*D*	*E*	
I	12.5	13.0	12.7	13.2	13.0	11.6	11.8	12.3	100.1
II	12.2	12.7	12.3	13.0	12.0	11.6	10.4	11.5	95.7
III	13.0	12.2	12.2	13.2	12.7	12.3	11.6	12.7	99.9
IV	12.3	13.0	12.8	12.0	12.0	12.5	11.4	12.0	98.0
V	11.9	12.7	12.8	12.8	13.2	11.8	11.4	12.5	99.1
VI	11.9	12.2	(*y*)	12 8	12.2	12.3	10.6	12.8	84.8
Variety Totals	212.4			77.0	75.1	72.1	67.2	73.8	577.6

ANSWER: The design of your experiment is not orthodox in that the X-plots are not randomized in the blocks. For this reason I would prefer to exclude these plots from the estimate of error. This error should be based on the five seedlings randomized in the six blocks.

I suspect that the average variance among the X-plots is biased downwards because these plots would seem to be less widely distributed than the others. However that may be, in this experiment the variance among them (0.135) is less than half the discrepance among the seedlings (0.286). If my fears are well founded, the error for the X-variety (unknown because of lack of randomization) is different from that of the seedlings. But the only way to make tests of significance is to assume that the error for the seedlings applies to the X-variety as well. So the estimate of error does not involve the missing plot.

If the X-plots were randomized along with the seedling plots, minimizing the appropriate experimental error would be accomplished by substituting the following value for the missing X-plot:

$$y = \frac{cbB + (t + c - 1)T - cG}{cb(t + c - 2) - t + 1},$$

where c is the number of X-plots per block, the other letters having the usual meanings. Despite its lack of validity, a numerical illustration is based on your data:

$$t = 6, \quad b = 6, \quad c = 3, \quad B = 84.8, \quad T = 212.4, \quad G = 577.6.$$

Substituting, $y = 12.34$.

As usual, this technique results in a treatment mean square which is slightly biased upward, but the bias is usually considered negligible.

In the cited experiment, conclusions will be the same regardless of the method used; no seedling significantly outranks the standard variety.

109 **QUERY:** We are conducting tests on the efficiency of several types of cotton pickers. The ultimate aim is to obtain the efficiency for each of these types of pickers, and to compare their efficiency under several field conditions.

Usual statistical methods of analysis of the data could be used if all pickers were operated simultaneously under each field condition. Physical limitation, however, prevent our doing this. The only procedure open to us now is to operate one picker only, picker A, for example, under all field conditions and to operate each of the other pickers on only one of these fields. To be more specific, using symbols:

Picker A and B can be used on field 1, pickers A and C can be used on field 2, pickers A and D can be used on field 3, etc.

Our problem then, is to find a statistically correct method, if one exists, to compare pickers A, B, C, etc. We would appreciate very much your advice on a method of analyzing the data which we collect from such a procedure. I should also say that adequate replications will be made in each field.

ANSWER: In each field you will have a replicated comparison of two pickers. I assume that they will be operated in pairs of plots (rows or swaths) resulting in a number of blocks with randomization of the positions of the pickers in each block. This will lead to the following analysis of variance:

Source of Variation	Degrees of Freedom	Mean Square
Blocks	$b - 1$	
Pickers	1	
Error	$b - 1$	s^2

The experiment will also provide estimates, $\bar{x}_A$ and $\bar{x}_B$, of the efficiencies of the two pickers, A and B. The significance of the difference between them will be tested in the usual manner. This procedure, repeated in other fields, will give the desired comparisons of A and C, D, etc.

The comparison of two pickers such as B and C involves some assumptions about the effect of field conditions on what may be called the *true* efficiency of the pickers A, B and C. If field conditions have *no* effect on the true efficiencies, which may be your "ultimate aim", then you can compare $\bar{x}_B$ and $\bar{x}_C$ directly. Assuming that the experimental errors, s_1^2 and s_2^2 in the two fields, are random samples from a common σ^2, their sums of squares can be combined in the usual fashion and the t-test applied.

You may find it more realistic to assume that field conditions affect the true efficiencies of two pickers by some additive constant characteristic of each sampled field. That is,

in field 1: $$A_1 = \pi_A + \phi_1 + e_{11}$$

$$B_1 = \pi_B + \phi_1 + e_{12} ,$$

whence: $$A_1 - B_1 = \pi_A - \pi_B + \text{errors},$$

where π is the true efficiency and ϕ_1 is the additive constant for field 1. Similarly, in field 2: $A_2 - C_2 = \pi_A - \pi_C +$ errors.

$$\text{Subtracting: } (A_2 - C_2) - (A_1 - B_1) = \pi_B - \pi_C + \text{errors.}$$

That is, subtraction of the two differences provides an estimate of the true difference between the efficiencies of B and C. Again assuming a common σ^2, this estimate has a mean square which is four times the pooled mean square of the two fields.

This is an unnecessarily expensive kind of experiment because: (i) the efficiency of picker A is evaluated with great precision at the expense of precision in the other pickers; (ii) the experiment is insensitive owing to the large mean square for the comparison of B, C, etc.; (iii) the comparisons are not all independent.

I suggest that you consider the balanced incomplete block experiment which you will find described and illustrated in Cochran and Cox "Experimental Designs", Chapter 11. Plans suitable for various numbers of pickers are laid out on pages 329-331. In this type of experiment all pairs of the pickers are tried, each in one of the fields, and all differences are evaluated with equal precision. As above, you will be assuming additive field effects; that is, no interaction of picker efficiencies with field conditions.

As your experience increases, you may learn that none of the above assumptions is suitable. If so, the design will have to be altered to comply with the newly found facts.

ABSTRACTS

Meeting of The Biometric Society, French Region, February 3, 1954

270 J. M. FAVERGE. **Un Exemple d'Adaptation de l'Analyse de la Variance a un Probleme Psychologique.**

"La méthode d'analyse de la variance a besoin d'être adaptée pour permettre d'exploiter les données expérimentales en psychologie. Ainsi, on rencontre fréquemment dans ce domaine des tableaux carrés de h lignes et h colonnes où la diagonale joue un rôle particulier et où l'on associe les cases symétriques par rapport à cette diagonale. C'est le cas où l'on recueille les jugements des h membres d'un groupe sur les membres du groupe; la diagonale contient les auto-jugements et les cases symétriques par rapport à la diagonale les jugements réciproques. C'est aussi souvent le cas en psychologie expérimentale, par exemple, dans les expériences du type de celles de Fitts et Seeger sur la compatabilité des stimuli et des réponses.

L'exposé avait pour premier objectif de montrer comment on peut extraire un degré de liberté du résidu afin de permettre la comparaison des termes diagonaux aux autres; cette comparaison est

$$d = (m_u - m_y)\sqrt{h - 1}$$

où m_u est la moyenne des h nombres diagonaux et m_y la moyenne des $h(h - 1)$ nombres non diagonaux.

Le deuxième objectif était de donner une méthode permettant d'étudier les jugements réciproques; elle est fondée sur la décomposition des $(h - 1)^2$ degrés de liberté du résidu en $[(h - 1)(h - 2)]/2$ degrés de liberté correspondant à la somme des carrés

$$\tfrac{1}{4} \sum (x_y - x_{ji} - m_i - m'_j + m_j + m'_i)^2$$

et en $[h(h - 1)]/2$ degrés de liberté correspondant à la somme des carrés

$$\tfrac{1}{4} \sum (x_y + x_{ji} - m_i - m'_j - m_j - m'_i + 2m)^2$$

où les m non accentués sont des moyennes de lignes et les m accentués des moyennes de colonnes".

271 J. M. LEGAY. **L'Aspect Biometrique dans l'Etude du Comportement Alimentaire Chez le Ver a Soie: Donnees sur l'Apprentissage dans la Recherche de la Nourriture.**

Les données expérimentales recueillies à ce sujet ont permis:
1—une étude du comportement moyen des Vers avec détermination de

l'allure générale du phénomène, comparaison des performances successives et de leur variance, ajustement de la courbe d'apprentissage à une courbe théorique, examen des variations selon l'âge des Vers.
2—une étude du comportement individuel des Vers, avec recherche de corrélation entre variables et entre rangs et détermination de types de comportement.

En conclusion, s'il est facile de caractériser de façon quantitative le comportement moyen des Vers, il est par contre difficile de prévoir d'après quelques essais la valeur des performances individuelles à la fin de l'apprentissage. Il aurait été intéressant de trouver des tests rapides permettant d'envisager une sélection.

272 G. E. P. BOX and S. L. ANDERSEN. (North Carolina State College.) **Effect of Non Normality and Variance Inequality on Statistical Tests.** (*By Title*)

Excerpts are made from a large number of tables in the literature which estimate the effect of failures of the assumptions on the type I error of several tests for comparing means and variances. New tables have been prepared for some of these tests, using approximate permutation tests to aid in the evaluation of the effect of non normality and variance inequality. For some cases these new tables are compared with results previously obtained by more complex means.

The effect of these failures of assumptions can also be expressed in several cases as a modification of the degrees of freedom of the standard tests.

An empirical sampling experiment has been performed to compare the type I error and power of normal-theory for tests variances with newer tests of the robust type in situations where the parent population is not normal.

273 ROBERT M. ABELSON and RALPH ALLAN BRADLEY. (Virginia Agricultural Experiment Station and Virginia Polytechnic Institute.) **A 2 x 2 Factorial with Paired Comparisons.**

The parameters previously specified for a method of paired comparisons are redefined in such a way as to permit the use of treatments in factorial array. The algebraic procedure is shown in general but the normal equations resulting from the use of maximum likelihood are nonlinear and difficult to solve. Easy solution of the normal equations

seems to be limited to the 2 x 2 factorial and an explicit solution is given for that case.

The method of paired comparisons presented for 2 x 2 factorial treatments permits most of the comparisons available through usual analysis of variance. It is possible to test for the presence of both main effects and their interaction.

A numerical example is included.

274 M. C. K. TWEEDIE. (Virginia Polytechnic Institute.) **Some Theorems on Unbiased Systems of Confidence Intervals.**

Taking an unbiased system to be one in which the true value of the parameter (θ) is at least as likely to be covered by the confidence interval as any other value (cf. *Ann. Math. Stat.*, Vol. 24 (1953), p. 139), and not restricting the observed variate to be continuous, proofs are given of some simple theorems concerning the probability $A(\theta_0 \mid \theta')$ that a value θ_0 will be covered when θ' is true. For example, if θ is one dimensional and $A(\theta_0 \mid \theta')$ is a differentiable function of θ' at all θ_0 in some continuous set, then $A(\theta \mid \theta)$ cannot have a discontinuous decrease in gradient as θ increases through that set.

275 W. A. THOMPSON, JR. (Virginia Polytechnic Institute.) **A Topic in Variance Components Analysis.**

A lemma is proved which may sometimes be used to find the class of all statistics whose distributions are independent of the nuisance parameters. The least squares model with errors arising from two sources is then discussed, and the lemma is then applied to this case. These results are then specialized to partially balanced incomplete block designs.

276 R. G. PETERSEN. (North Carolina State College.) **The Distribution of Excreta by Freely Grazing Animals and its Effect on Pasture Fertility.**

The relative frequency of occurrence of 10 x 10 ft. squares containing 0, 1, 2, $\cdots$ excreta per square was determined for several small and one large pasture. The empirical distribution thus obtained was compared with several theoretical distribution functions, such as the Poisson and negative binomial distributions, which might be used to represent pastures in general.

The time at which each deposition occurred was combined with estimates of the rate of application of certain fertilizer elements, and with

functions describing the rate of loss of these elements from the root zone of the soil to obtain the probability distribution of fertility levels in the pasture. The empirical excreta distribution and the simpler theoretical distributions were compared to determine the general applicability of these simple functions in predicting the effect of excretal return on pasture fertility.

The results indicate that in determining the probability distribution of fertility levels it may safely be assumed that excreta are deposited in a Poisson fashion.

ARNOLD H. E. GRANDAGE. (North Carolina State College.)

277 Biological Assay of a Material when Interfering Substances are Present.

In the bioassay of a mixture, the observed responses may be postulated to obey the following model,

$$Y = \beta_0 + \beta_1 \log (X_1 + KX_2) + \beta_2(X_2) + \epsilon$$

where Y is the response metameter, X_2 is the dose of the mixture and X_1 is an added dose of a pure preparation of the component of the mixture that is to be determined. The proportion of this component in the mixture is K and the problem is to form an estimate for K.

Various designs were studied by use of empirical samples from known populations. Least squares estimates of K were computed using successive approximations to K until a minimum residual sum of squares was obtained. Confidence limits for K were computed by a "sliding sum of squares" method.

These empirical results were compared with the known parameter values and with asymptotic values. In general, the point estimates of K were biased, but the confidence limits were quite good.

S. M. FREE. (North Carolina State College.)

278 Relationships of Color Measurements and Some Quality Indices of Flue Cured Tobacco.

Optical instrument color ratings that define color by three continuous parameters (Brightness, Yellowness and Red to Green) were taken on samples of flue cured tobacco. These color measurements were related to market price by four linear models. The models range in complexity from a simple function of only the color indices to a function considering all parts of the government grade. The utility of the color measure-

ments and the effect of the models is determined for two different samples.

In addition, canonical correlations were determined to relate the instrument readings to the government color designators.

279 J. S. HUNTER. (North Carolina State College.) **Some Third Order Composite Designs.**

In attempting to estimate an unknown continuous response function $\eta = F(X_1, X_2, \cdots, X_k)$ where η is the response variable and $X_1, X_2, \cdots . X_k$ are quantitative independent variables, it is assumed possible to replace the function by its Taylor's Series. The coefficients of the Taylor's Series approximation of the unknown function may then be estimated by least squares. Recently, the construction and use of composite designs for the purpose of fitting these second order models has stimulated considerable interest. However, situations arise in which lack of fit of the second order approximation requires the estimation of the coefficients of a third order model. Furthermore, the failure to estimate third order effects may affect the estimates of terms of lower order. Some third order composite designs are discussed and their application to a problem in chemical engineering demonstrated.

280 U. KRECH and D. KODLIN. (University of Pittsburgh.) **The Bioassay of Poliomyelitis Vaccines in Mice.**

Quantal and quantitative response data are available for the evaluation of relative potency of polio vaccines in mice. Probit–log dose and log antibody–log dose metameters are satisfactory. Though there is indication of dissimilar mode of action for preparations produced by different methods, so far no dissimilarity could be demonstrated within methods. The inherent precision of the test is of the order of 0.3 for quantitative and 0.8 for quantal response types.

281 HAROLD F. HUDDLESTON. (Federal-State Crop Reporting Service, Raleigh, N. C.) **Generalized Regressions for Weather Factors.**

The inverse matrix approach is used to determine "Gauss Multipliers" primarily to reduce the amount of the computations when the same set of independent variables is used for a number of crops or dependent variables. The stability of the parameters for the Gauss Multipliers over time for certain combinations of monthly weather factors for a

fairly large homogeneous area is investigated. When these parameters stabilize, the use of lengthy weather records is preferable to the use of weather data for shorter periods corresponding to some sub-period for which individual crop data may be available. The covariance terms between crop yields (dependent variables) and the weather factors are determined for only the sub-period corresponding to the yield data and are used with the covariance terms or Gauss Multipliers relating to the independent variables for the longer period of record.

The use of a general set of Gauss Multipliers or "population values" appears possible as indicated by the preliminary analysis, but dependent upon: (1) Finding a quick method of estimating a factor of proportionality, K, by which one can convert to the true units or coefficients, or (2) using the ratios of the Cij's (elements of the inverse matrix) to compute regression coefficients proportional to the net regression coefficients.

282 GEORGE KARREMAN. (The University of Chicago.) **The Resonance of the Arterial System.**

The arterial system is assumed to consist of two elastic chambers connected by a conducting channel. It is assumed that a current of fluid enters one chamber, whereas the other chamber is drained by a pipe with a certain peripheral resistance. The continuity of the fluid is described by a differential equation for each chamber. The inertia resistance of the conducting channel is taken into consideration.

It is shown that the system may possess a resonance frequency. The latter, if it exists, as well as the damping coefficients are expressed in terms of the elastic moduli of the chambers, the conductivity of the channel, and the peripheral resistance. It is shown that with plausible values of the latter variables the resonance frequency as determined theoretically has the right order of magnitude as found experimentally.

283 FRED H. HULL. (Florida Agricultural Experiment Station.) **Multigenic Population Models with No Algebra and No Statistic beyond the Arithmetic Mean.**

Many students in genetics courses with little or no functional knowledge of algebra, or statistics more than a simple average, need population multigenics presented objectively in understandable terms. Similar presentation to mathematical statisticians avoiding inhibitions of intuitive obsessions of present day genetics lore, may set the stage for solution of some of the more intricate problems.

First Biometric Colloquy of the German Section of The Biometric Society Bad Nauheim (Kerckhoff-Institute), January 15–17, 1954

284 H. GEBELEIN, Bamberg. **Three Types of Statistical Inferences.**

(1) The title refers to the inferences from a sample to a sub-sample, from a sample to a finite population which contains the sample, and finally from one sample to another sample of the same finite population. These inferences are treated in detail in the book "Zahl und Wirklichkeit", following a suggestion by Wagemann. The conclusions from one finite set to another finite set are reached without an excursion into infinity. Purpose and advantage of this finite reasoning.

(2) Mutual connections among the three inferences. Respective comparison with the hypergeometric distribution, Bayes' distribution, and Greenwood's result.

(3) Symmetries and group characteristics revealed by the mathematical equations which describe the three inferences. Tentative formulation of the special problems involved in the inference from a sample to an enclosing population.

(4) General laws for the three inferences if they are applied to k different attributes. Their changes—distinct from each other—if this number k is reduced. A strange equation of equivalence. An open problem.

285 H. GEIDEL, Rethmar. **Mathematical Fundamentals of the Analysis of Variance and the Design of Experiments.**

The least square method by Gauss. Analysis of variance. Different types of this method. Separation of variances. F-test, t-test. Connection between these two tests.

286 H. W. VON GUÉRARD, Duesseldorf. **Biometrical Statistics Suggesting a Structure.**

Biometrical populations used to be treated according to purely mathematical methods, like those developed by actuaries, in which no special assumptions are involved. With this respect the possibility may be mentioned to choose parameters of a power series or an exponential expression so that few terms interpolate the observed mortality. Lexis started to explain a given mortality as a sum of three distributions —not necessarily normal—the mortality of the infants, untimely deaths

at the height of life, and the mortality in old age. This separation according to prevailing causes introduces into the purely arithmetic scheme a structure which indicates a logical connection between the phenomenon and the interpolating formula.

This is a step from pure interpolation to an explanation. The equations which describe numerically the distribution are more than a mechanism built to throw out a set of reliable figures for the purpose of predications. They form a one-to-one correspondence between the observed and the mathematical distribution.

The following further examples are mentioned: life expectancies of married and single men; studies on intervals between births of children to the same parents. It is believed that the parameters of such structural formulae, especially the variability of these parameters to changed conditions, suggest an approach to a causal analysis. It is expected that these formulae remain reliable even at a higher variability because of the genuine fit of the curves as opposite to a purely interpolating equation. It is an open problem whether a looser interpretation of confidence intervals should be permitted if structural formulae are applied.

287 F. KEITER, Hamburg. **Statistical Treatment of Compounded Attributes in Proving the Paternity by Using Anthropological and Hereditary Traits.**

The proof of paternity, based on the similarity of many traits, is in its core a purely statistical method although empirical knowledge and vague estimations are involved frequently in practical cases. Simple one-dimensional attributes are assumed for the statistical treatment recommended by Essen-Moeller and Keiter. However compounded similarities are striking in many cases, for instance in the face. They use to be exploited by the empirics. They may be included into a strict evaluation too if the scores of similarity do not refer to the single traits but to the whole pattern, for instance the frontal region, view of the nasal area from below, shape of the pinna, etc. Recent successes, gained by this method, are reported.

288 A. LEIN, Schnega, Hannover. **Application of Fisher's Methods to the Design and the Performance of Agricultural Experiments.**

(1.) Principles for the design and the evaluation of trials.
(2.a) Particular problems involved in agricultural and horticultural experiments.
b) Structure of a simple, normal experiment.

c) Evaluation of an experiment by using the analysis of variance.
d) Control experiment. Example.
e) Consequences for size and shape of the plots.
f) Effect of the number of repetitions.
(3.) Further possibilities of designing.
a) Orthogonal schemes.
b) Non-orthogonal arrangements
c) Remarks on the efficiency of the designs.
(4.) Reference to other methods applied to agricultural experiments.

289 H. MUENZNER, Goettingen. **Problems and Conclusions of Mathematical Statistics.**

It is shown that mathematical and statistical methods are adequate and even necessary in all branches of research. Specific models of the mathematical statistics are discussed. They are needed for explaining the results which may be gained by the application of statistical methods. The main principles are surveyed on which estimations and tests depend. Differences of approach and opinion are discussed.

Special fields of mathematical statistics are mentioned because of their practical importance, namely the analysis of variance, factor analysis, separation of distributions.

Finally it is indicated how the mathematical statistics developed from the evaluation of given data to the design of experiments and to sequential analysis. It is now a tool of research which is needed at all stages of scientific work.

290 W. SIECKMANN, Steinhude. **Determining Curves of Reactions by Using Probits and Logits.**

If it is studied how the ratio of the responding to the exposed test items depends on the concentration of a poison, a convenient function with two degrees of freedom used to be assumed as curve of reaction. The two parameters of the function have to be estimated according to the observations. If the parameters refer to the localization of the mean and to the variance, workable estimates can be found for all types of functions. It is the peculiarity of the method in question that the curve of reaction is transformed into a straight line before the parameters are estimated. Thus the problem is reduced to the estimation of the two parameters of a straight line. Functions most frequently applied are the normal distribution and the logistic function. In the literature the corresponding methods are called probit respectively

logit analysis. They are surveyed and generalizations are discussed. They are needed in order to consider the natural mortality during the experimental period or an eventual immunity of the test animals.

291 E. WALTER, Goettingen. **Exhaustion of a Given Significance Level for Combinative Tests.**

Since the distribution of the tested variable x is discrete for combinative tests, in general it is impossible to choose a critical region which corresponds exactly to a given significance level. Therefore a region $x \geq x_a$ is used to which a probability $\bar{\alpha} \leq \alpha$ belongs. But a value $\bar{\bar{\alpha}} > \alpha$ would arise by including the adjacent point $x_b < x_a$ into the critical region. Since the significance level cannot be exhausted, the efficiency of combinative tests is lower than for tests working with a continuous variable. Without changing the essence of the test, it is possible to increase its efficiency. If x_b is observed, a further test is applied which uses the variable y. The hypothesis is refused if $y \geq y_a$. Here y_a is defined by

$$\int_{y_a}^{\infty} f(y \mid x_b)\, dy \leq \frac{\alpha - \bar{\alpha}}{\bar{\bar{\alpha}} - \bar{\alpha}} .$$

In the cases $x \geq x_a$ and $x < x_b$ the original test is applied according to the usual rule.

292 R. WETTE, Heidelberg. **The Sequential Probability Ratio Test.**

The best sample size N which is frequently known from preceeding experiments, at least approximately, is kept constant for ordinary statistical procedures, by which hypotheses are tested in biometrics. The size of the 'critical region', i.e. the probability of refusing incorrectly the hypothesis, may be stated arbitrarily in advance. The shape of the critical region is determined so that its potency, i.e. the complimentary probability for incorrectly accepting the hypothesis, becomes a maximum. The relative efficiency of this method is rather small. On the other hand, if size, potency, and best shape of the critical region is given, the sample size n becomes variable. A higher relative efficiency goes together with a decrease of the expected size of the sample which reaches in practical cases frequently about 50%. Starting from these ideas, the sequential probability ratio test was developed (Wald, Friedmann and Wallis, et al.), at first for industrial purposes. Using this method, the sample size is increased step by step

until the increasing weight is sufficient for making a decision in whatever direction. Minimum, maximum, and expectation of the sample size can be determined. A graph shows the potency of the procedure as function of the parameters of the population. The method is available in workable form for several problems which might be applied to biometrics. The numerical work is slight and in many cases it is reduced further by tables.

293 R. K. BAUER, Munich. **Discriminant Analysis.**

The discriminant analysis is the first completely abstract method by which different populations are separated. Several traits of the subjects are listed. It is possible to assign the subjects to the correct population according to their pattern of these traits. The 'Linear Discriminant Analysis' by R. A. Fisher (1936) which makes use of a linear combination of the traits of a subject, is the most convenient method from a numerical viewpoint, but it assumes normal distributions of the populations in question. The 'Quadratic Discriminant Analysis' by B. L. Welch (1939) avoids this assumption. In principle it is an optimum, but it is not workable. In order to simplify the calculations L. S. Penrose (1945) built from the different traits two statistics to which he applied Fisher's analysis. C. A. B. Smith (1947) transferred Penrose's statistics to Welch's analysis. Main fields open to a discriminant analysis are anthropology, psychology, and a quantification of qualitative attributes.

294 H. DOERING, Goettingen. **Calculation of the Hereditary Component of the Variance of Attributes (So-called Heredibility).**

The hereditary component of the variance of attributes is discussed and its importance for animal husbandry is emphasized. The computation of various estimates is presented by using a hierarchical model of the analysis of variance. The causes for differences between the proposed estimates are discussed. The calculation of the sampling variance of hereditability estimates is sketched.

295 E. WELTE, Bonn. **Design of Experiments in Clinical Medicine.**

It is shown that there are differences between experiments in science or biology and those in clinical medicine. The clinical experiment has to account for the fact that a sick human being is the subject of the

research. Control samples are possible only in the case of acute diseases (infections, poisoning). In the case of chronic illnesses the observation of different intervals during the sickness of the same patient (before, during, and after medication) substitutes for the comparison of two different groups of patients. Another peculiarity of the clinical experiment is the great number of co-operating factors. As far as possible, they have to be avoided.

Meeting of The Biometric Society, French Region, May 5, 1954

296 SULLY C. LEDERMANN, (Institut National d'Etudes démographiques). **La Mortalité par Causes Dans ses Rapports Avec l'Alcoolisation de la Population.**

L'alcoolisation excessive d'une population peut affecter de façon importante sa mortalité. L'étude des incidences de l'alcoolisation excessive de la population française adulte sur sa mortalité a été conduite en utilisant la Statistique des causes de décès. Une méthode a été mise au point, pour connaître la répartition par grandes causes (tuberculose, cancer, etc. . .) des décès dont la cause n'est pas spécifiée ou est mal définie. Cette méthode est applicable à des pays autres que la France.

Les taux de mortalité par causes, ainsi améliorés, ont permis de poursuivre les recherches de deux façons : 1° / en comparant l'évolution, dans le temps, de la mortalité pour certaines causes et de la surmortalité masculine, à celle de la consommation de vin et d'alcools; 2°/en analysant les corrélations présentées entre elles par les différentes causes de décès, dans les 90 départments français. Cette analyse a été effectuée selon les principes de l'analyse factorielle, telle qu'elle est employée par les psychotechniciens.

Les résultats obtenus forment un ensemble homogène : l'alcoolisation excessive paraît jouer un rôle important après 35 ans, notamment dans l'étiologie de la tuberculose pulmonaire, et probablement aussi dans celle de certains cancers. L'étude a montre, en outre qu'en France, la surmortalité masculine est étroitement liée, depuis un siècle, avec l'alcoolisation excessive des hommes.

297 D. BARGETON. **Interpretation de l'Action des Antithyroidiens sur le Metabolisme Basal.**

On peut prévoir une évolution du métabolisme en fonction exponentielle du temps par administration d'un antithyroïdien si :

a) le métabolisme est fonction linéaire de la quantité d'hormone thyroïdienne présente;

b) l'hormone disparaît à une vitesse proportionnelle à sa concentration;

c) l'antithyroïdien provoque d'emblée une réduction fixe de la production d'hormone.

L'observation de rats traités par différents antithyroïdiens fournit des données en accord avec ces hypothèses et donne une mesure de la vitesse de sécrétion de l'hormone thyroïdienne.

Si l'on prend comme réponse l'abaissement du métabolisme correspondant au niveau final d'équilibre, on obtient un diagramme linéaire probit de la réponse—log dose exprimant l'activité en pourcentage d'inhibition secrétoire.

Ces diagrammes permettent la comparaison d'activité de différents antithyroïdiens par les méthodes usuelles de standardisation.

THE BIOMETRIC SOCIETY

ENAR. The Region met on the campus of the University of Florida in Gainesville on March 18 and 19, 1954. At the opening session, held jointly with the Institute of Mathematical Statistics, papers on Truncation Problems and Applications were presented by A. C. Cohen, J. R. Duffett and John Woodward, with D. E. South as chairman. G. W. Snedecor officiated as chairman of the first afternoon session on Quantitative Genetics, with papers by Virgil Anderson and C. Clark Cockerham. Herbert A. Meyer chaired the following session on the Training of Statisticians in the South, discussed by G. E. Nicholson, Jr. and by Ralph A. Bradley. Two sessions of contributed papers completed the day's program, with G. L. Edgett and P. N. Somerville presiding. Abstracts of these papers will be printed in the *Annals of Mathematical Statistics.* On March 19 R. A. Bradley presided at the opening session of four contributed papers. Gertrude Cox then introduced M. G. Kendall who addressed the Region on "Biological Applications of Multivariate Analysis Techniques". Lee Crump presided at the opening afternoon session with three invited papers on procedures for multiple comparisons: "Multiple Range and Multiple *F* Tests" by D. B. Duncan; "Confidence Precedures are Better" by John W. Tukey; and "Some Applications of the Multiple Comparisons Tests" by R. J. Hader. W. F. Callander took the chair for a final session of four more contributed papers.

A joint evening symposium on "Biometric Methods in Immunology" was sponsored by the American Association of Immunologists and The Biometric Society (ENAR) before the Federated Societies on April 14 in Atlantic City, New Jersey. Dr. H. C. Batson, University of Illinois College of Medicine, served as chairman of a two-hour program of four papers as follows: (a) Official Standards for Immunology—A Challenge to Biometry. Lloyd C. Miller, Chairman of Revision, U.S. Pharmacopoeia, New York; (b) Problems in the Measurement of Immunity and of the Potency of Immunizing Agents. A. A. Miles, Director, The Lister Institute of Preventive Medicine, London, England; (c) The Practical Value of Sound Methods of Biological Assay. C. A. Morrell and Louis Greenberg, Food and Drug Divisions and Laboratory of Hygiene, Department of Natural Health and Welfare, Ottawa, Canada;

and (d) Is There an Increased Risk? Irwin Bross, Department of Public Health and Preventive Medicine, Cornell University Medical College, New York. Following the formal presentation of papers a lively and extensive discussion extended the meeting by more than one hour. Nearly 300 individuals attended at least part of the session and approximately 50 to 60 remained until the close of the discussion.

Région pour la Belgique et le Congo Belge. Une conférence était donnée lundi 26 avril à l'Institut d'Hygiene et de Médecine Sociale à Bruxelles par le Professeur P. Mahalanobis sur le sujet: Statistical Sampling. Une discussion suivait l'exposé du Professeur Mahalanobis.

Une réunion de la Société Adolphe Quetelet avait lieu mercredi 16 juin dans les locaux de la Fondation Universitaire à Bruxelles. Le colloquim était consacré à l'Agronomie et centré sur les problèmes de la betterave dans ses rapports avec la Biométrie. Programme: (1) Introduction, par Mr. L. Martin; (2) Exposé sur le problème de l'expérimentation des variétés de betterave et ses relations avec la Biométrie, par Mr. N. Roussel; (3) Exposé sur les quelques applications de la Biométrie, aux essais sur betterave (expérimentation des engrais, méchanisation des travaux de printemps, etc.), par M. R. Wauthy; (4) Discussion.

Région Française. Une réunion de la Société avait lieu le mercredi 5 Mai au Laboratoire de Zoologie de l'Ecole Normale Superieure. Ordre du Jour: S. Ledermann, La mortalité dans ses rapports avec l'alcoolisation de la population; D. Bargeton, Interprétation de l'action des médicaments anti-thyroidiens.